Communications in
**Mathematical
Physics**

Generating Diffusions with Fractional Brownian Motion

Martin Hairer[iD]**, Xue-Mei Li**[iD]

Imperial College London, London, UK. E-mail: m.hairer@imperial.ac.uk; xue-mei.li@imperial.ac.uk

Abstract: We study fast/slow systems driven by a fractional Brownian motion B with Hurst parameter $H \in (\frac{1}{3}, 1]$. Surprisingly, the slow dynamic converges on suitable timescales to a limiting Markov process and we describe its generator. More precisely, if Y^ε denotes a Markov process with sufficiently good mixing properties evolving on a fast timescale $\varepsilon \ll 1$, the solutions of the equation

$$dX^\varepsilon = \varepsilon^{\frac{1}{2}-H} F(X^\varepsilon, Y^\varepsilon)\, dB + F_0(X^\varepsilon, Y^\varepsilon)\, dt$$

converge to a regular diffusion without having to assume that F averages to 0, provided that $H < \frac{1}{2}$. For $H > \frac{1}{2}$, a similar result holds, but this time it does require F to average to 0. We also prove that the n-point motions converge to those of a Kunita type SDE. One nice interpretation of this result is that it provides a continuous interpolation between the time homogenisation theorem for random ODEs with rapidly oscillating right-hand sides ($H = 1$) and the averaging of diffusion processes ($H = \frac{1}{2}$).

Contents

1. Introduction

The setting considered in this article is as follows. Consider a particle in a rapidly evolving random medium, so that it is governed by a stochastic differential equation of the type $dx_t = A(x_t, t/\varepsilon)\,dt + \sigma(x_t, t/\varepsilon)\,dB$[1] for a small parameter $\varepsilon > 0$. The situation we are interested in is where, in the "static" case (i.e. when A and σ have no explicit time dependence), the system is either super- or subdiffusive. This is the case if the driving noise B is modelled by fractional Brownian motion (fBM) with Hurst parameter $H \neq \frac{1}{2}$. Recall that fractional noises (i.e. the time derivative of fBM) can be obtained as scaling limits in statistical mechanics models [11,38,65] and that fBM with Hurst parameter H is a Gaussian process with stationary increments and self-similarity exponent H. It is therefore characterised (up to an irrelevant global shift) by the fact that $\mathbf{E}(B_t - B_s)^2 = |t - s|^{2H}$, so that it is superdiffusive for $H > \frac{1}{2}$ and subdiffusive for $H < \frac{1}{2}$. The covariance of its increments, $\mathbf{E}(B_{t+1} - B_t)(B_{s+1} - B_s)$, decays at rate $|t - s|^{2H-2}$ for large $|t - s|$ and therefore exhibits long-range dependence when $H > \frac{1}{2}$.

We furthermore assume that the rapid time evolution of the environment is described by a hidden Markov variable, thus leading to the model

$$x_t^\varepsilon = x_0 + \int_0^t F(x_s^\varepsilon, y_s^\varepsilon)\,dB_s + \int_0^t F_0(x_s^\varepsilon, y_s^\varepsilon)\,ds, \qquad (1.1)$$

with B an fBM with Hurst parameter $H \in (0, 1)$ in $\mathbf{R}^m$ and $F(x, y) \in \mathbb{L}(\mathbf{R}^m, \mathbf{R}^d)$. The stochastic integral appearing in the first term is problematic when $H < \frac{1}{2}$: one should really interpret this equation as $x_t^\varepsilon = \lim_{\delta \to 0} x_t^{\varepsilon, \delta}$ with $x_t^{\varepsilon, \delta}$ driven by a smooth approximation B^δ to B with relevant timescale $\delta \ll \varepsilon \ll 1$, see Sect. 1.1 below. Regarding the fast Markov variable, a prototypical situation is that of a system of the type

$$dy_t^\varepsilon = \sigma(x_t^\varepsilon, y_t^\varepsilon)\frac{dW_t}{\sqrt{\varepsilon}} + b(x_t^\varepsilon, y_t^\varepsilon)\frac{dt}{\varepsilon}, \qquad (1.2)$$

where W is a Wiener processes independent of the fBM B appearing in (1.1). This allows for the case where the variable x feeds back into the evolution of y, but for most of this article we assume that there is no x-dependence in (1.2). We also assume that y_t admits a unique invariant probability measure μ. In the case with feedback, we have a family of invariant measure μ_x obtained by "freezing" the value of the variable x in (1.2).

---

[1] Following the tradition of probability theory, the subscript t denotes dependence on the time parameter, not differentiation.

It was recently shown by the authors in [31] that in the case $H > \frac{1}{2}$ the process x_ε converges *in probability* to the solution to

$$dx = \bar{F}(x)\,dB + \bar{F}_0(x)\,dt, \tag{1.3}$$

where the average of any function h is given by $\bar{h}(x) = \int h(x, y)\,\mu_x(dy)$. The aim of the present article is to investigate the two cases left out by the aforementioned analysis, namely what happens when either $H < \frac{1}{2}$ or when $H > \frac{1}{2}$ but $\bar{F} = 0$ in (1.3)?

1.1. Description of the model. It turns out that the effect of the rapid oscillatory motion described by the fast variable y is to slow down the motion of x in the superdiffusive case and to speed it up in the subdiffusive case. This can be explained by the following heuristics. For times of order $t \lesssim \varepsilon$, the process Y doesn't evolve much so that, by the scaling property of the driving fBM, one expects the process x to move by about ε^H in a time of order ε. On large times $t \gtrsim \varepsilon$ on the other hand we will see that the limiting process is actually Markovian, even in the case with long-range dependence. This suggests that over times of order t the process x performs about t/ε steps of a random walk with step size ε^H and therefore moves by about $\varepsilon^H \sqrt{t/\varepsilon}$. This suggests that one should multiply F by $\varepsilon^{\frac{1}{2}-H}$ in order to obtain a non-trivial limit.

As a consequence, the equations we actually study in this article are of the form:

$$dX^\varepsilon = \varepsilon^{\frac{1}{2}-H} F_i(X^\varepsilon, Y^\varepsilon)\,dB_i + F_0(X^\varepsilon, Y^\varepsilon)\,dt, \qquad Y^\varepsilon(t) = Y(t/\varepsilon), \tag{1.4}$$

(summation over i is implied), where B is a fractional Brownian motion with Hurst parameter H ranging from $\frac{1}{3}$ to 1^2, and Y is an independent stationary Markov process with values in some Polish space $\mathcal{Y}$, invariant measure μ and generator $-\mathcal{L}^3$. At the moment, we are unfortunately unable to cover the case when X feeds back into the dynamics of Y. When $H > \frac{1}{2}$, we furthermore assume that $\int F_i(x, y)\,\mu(dy) = 0$ for every $i \neq 0$ and every x.

Our main result is that, as $\varepsilon \to 0$, solutions to (1.4) converge in law to a limiting Markov process and we provide an expression for its generator. In fact, we have an even stronger form of convergence, namely we show that the *flow* generated by (1.4) converges to the one generated by a limiting stochastic differential equation of Kunita type (i.e. driven by an infinite-dimensional noise).

Remark 1.1. Of course, (1.4) is not quite the same as (1.2) which was our starting point. One way of relating them more directly is to perform a time change and set $X^\varepsilon = x_\varepsilon(\varepsilon^{1-2H} t)$ with x_ε solving (1.2). Then X^ε solves the equation $dX^\varepsilon = \tilde{\varepsilon}^{\frac{1}{2}-H} F_i(X^\varepsilon, Y^{\tilde{\varepsilon}})\,dB_i + \tilde{\varepsilon}^{\frac{1}{2H}-1} F_0(X^\varepsilon, Y^{\tilde{\varepsilon}})\,dt$ where we have set $\tilde{\varepsilon} = \varepsilon^{2H}$. When $H < \frac{1}{2}$, this then converges to the same limit as (1.4), but of course with F_0 in (1.4) set to 0. When $H > \frac{1}{2}$, then one would need to take F_i centred in (1.2) in order to obtain a non-trivial limit and our results imply that one again converges to the same limit as (1.4), at least in the case $F_0 = 0$.

---

[2] We could probably deal with $H \in (\frac{1}{4}, \frac{1}{3}]$ with our techniques, but this would obscure some of the arguments for relatively little gain. For $H \leq \frac{1}{4}$, there exists no solution theory even in the absence of Y.

[3] The convention of adding a minus sign to the generator simplifies our expressions later on.

The special case when $F_0 = 0$ and the F_i are independent of the x-variable yields a functional central limit theorem for stochastic integrals against fractional Brownian motion. This already appears to be new by itself and might be of independent interest.

As already hinted at, the map $t \mapsto F_i(\cdot, Y_t^\varepsilon)$ is too irregular to fit into the standard theory of differential equations driven by a fractional Brownian motion, especially when $H < \frac{1}{2}$, so that it is not even completely clear a priori how to interpret (1.4) for fixed $\varepsilon > 0$. These questions will be addressed in more detail in Sect. 2 below. Let us put these aside for the moment and consider the following ordinary differential equation

$$\dot{X}_s^{\varepsilon,\delta} = \frac{\delta^{H-1}}{\sqrt{\varepsilon}} v\left(\frac{s}{\varepsilon\delta}\right) F(X_s^{\varepsilon,\delta}, Y_s^\varepsilon) + F_0(X_s^{\varepsilon,\delta}, Y_s^\varepsilon), \tag{1.5}$$

where v is a smooth stationary Gaussian random process with covariance C such that $C(t) \sim |t|^{2H-2}$ for $|t|$ large. When $H < \frac{1}{2}$ we furthermore assume that $\int C(t)\,dt = 0$ and, when $H = \frac{1}{2}$, we assume that C decays exponentially and satisfies $\int C(t)\,dt = 1$. One way of obtaining such a process v is to set $v = \phi * \dot{B}$ for ϕ a Schwartz test function integrating to 1 (and $*$ denoting convolution in time). This in particular shows that, at least in law, one has $(\varepsilon\delta)^{H-1} v\left(\frac{t}{\varepsilon\delta}\right) = (\phi_{\varepsilon\delta} * \dot{B})(t)$, where we set $\phi_\varepsilon(t) = \varepsilon^{-1}\phi(t/\varepsilon)$. Since this converges in law to $\dot{B}$ as $\varepsilon\delta \to 0$, we can view (1.5) as an approximation to (1.4).

It is then possible to show that the limit $X^\varepsilon = \lim_{\delta \to 0} X^{\varepsilon,\delta}$ exists and our results hold with X^ε interpreted in this way. Furthermore, we will see that all our results hold uniformly over $\delta \in (0, 1]$ as $\varepsilon \to 0$. This in particular shows that the converse limit obtained by first sending $\varepsilon \to 0$ and then $\delta \to 0$ is the same, as are all limits obtained by other ways of jointly sending $\varepsilon, \delta \to 0$.

1.2. Description of the main results. We now give a precise formulation of our main results, albeit with a simplified set of assumptions. The reason is that while the simplified assumptions are straightforward to state, they are very stringent regarding the Markov process Y. The more realistic set of assumptions used in the remainder of the article however is quite technical to formulate. We first recall the following standard definition of the fractional powers of the generator of the process Y.

Definition 1.2. We write $\mathcal{H} = L^2(\mu)$ with μ the invariant measure of Y and $\langle \cdot, \cdot \rangle_\mu$ for its scalar product. For $\alpha \in (0, 1)$, we then say that $f \in \text{Dom}(\mathcal{L}^\alpha)$ if, for every $g \in \mathcal{H}$, the integral

$$\frac{1}{\Gamma(-\alpha)} \int_0^\infty t^{-\alpha-1} \langle P_t f - f, g \rangle_\mu dt < \infty,$$

converges and determines a bounded functional on $\mathcal{H}$ (which we then call $\mathcal{L}^\alpha f$). Recall that the generator of the process Y is $-\mathcal{L}$, so that $\mathcal{L}$ is indeed a positive operator in the reversible case and $\mathcal{L}^\alpha$ does then coincide with the definition using functional calculus.

Similarly, for $\alpha \in (-1, 0)$, we write $\mathcal{L}^\alpha$ for the operator given by

$$\mathcal{L}^\alpha f = \frac{1}{\Gamma(-\alpha)} \int_0^\infty t^{-\alpha-1} P_t f\,dt.$$

Since $t \mapsto t^{-\alpha-1}$ is locally integrable, it follows from the first point of Assumption 1.3 below that $\mathcal{L}^\alpha$ is a bounded operator on the subspace of $\text{Lip}(\mathcal{Y})$ consisting of mean zero functions.

Assuming that X^ε takes values in $\mathbf{R}^d$, we then define the $d \times d$ matrix-valued function

$$\Sigma(x, \bar{x}) = \frac{1}{2}\Gamma(2H+1)\sum_{k=1}^{m}\int F_k(x, y) \otimes \left(\mathcal{L}^{1-2H}F_k\right)(\bar{x}, y)\,\mu(dy), \qquad (1.6)$$

where $\mathcal{L}$ acts on the second argument of F_k. As we will see in Remark 1.7, the expression (1.6) is naturally interpreted as the limit $\delta \to 0$ of a "local" Green–Kubo formula associated to the fluctuations of (1.5).

Note that the condition $\bar{F}_k = 0$ is necessary in the case $H > \frac{1}{2}$ since the negative power of $\mathcal{L}$ appearing in this expression does not make sense otherwise, see also Remark 2.5 below. We shall assume mixing conditions and Hölder continuity of the Y variable, see Assumptions 2.1–2.3 below, as well as a regularity condition on $x \mapsto F(x, \cdot)$ (and also F_0) as spelled out in Assumption 2.7. A simpler set of conditions is as follows, the first of which is a strengthening of Assumptions 2.1 and 2.3, the second is a strengthening of Assumption 2.2, and the last is just a restatement of Assumption 2.7 in this context.

Assumption 1.3 *(Simplified Assumptions). The functions F_i appearing in (1.4) as well as the Markov process Y satisfy the following.*

1. *The Markov semigroup associated to the process Y is strongly continuous and has a spectral gap in $\mathrm{Lip}(\mathcal{Y})$, the space of bounded Lipschitz continuous functions on $\mathcal{Y}$.*
2. *In the case $H < 1/2$ we assume that, for any $\alpha < H$, the process $t \mapsto Y_t$ admits α-Hölder continuous trajectories and its Hölder seminorm (over intervals of length 1 say) has bounded moments of all orders.*
3. *When $H > \frac{1}{2}$, we also assume that $\int F_i(x, y)\,\mu(dy) = 0$ for every $i \neq 0$ and every x.*
4. *There exists $\kappa > 0$ such that, for every $i \geq 0$, $x \mapsto F_i(x, \cdot)$ is $\mathcal{C}^4$ with values in $\mathrm{Lip}(\mathcal{Y})$ and its derivatives of order at most 4 are bounded by $C(1+|x|)^{-\kappa}$ for some $C > 0$.*

Remark 1.4. Recall that a Markov semigroup $(P_t)_{t\geq 0}$ admits a spectral gap in any given Banach space $E \subset L^2(\mu)$ if $P_t \colon E \to E$ is a bounded linear operator for every t and if there exist constants $c, C > 0$ such that $\|P_t f - \mu(f)\|_E \leq Ce^{-ct}\|f\|_E$ for all $f \in E$. For this definition to make sense, E of course needs to contain all constant functions.

The reason why we are aiming for a more general result at the expense of a much more technical set of assumptions is that having a spectral gap in $\mathrm{Lip}(\mathcal{Y})$ is a very restrictive condition which is not even satisfied for the Ornstein–Uhlenbeck process.[4]

Theorem 1.5. *Let $H \in (\frac{1}{3}, 1)$ and let Assumption 1.3 hold. For fixed $\varepsilon > 0$, $\alpha < H$, and $T > 0$, the process $X^{\varepsilon,\delta}$ converges in law in $\mathcal{C}^\alpha([0, T])$ as $\delta \to 0$ to a limit X^ε which we interpret as the solution to (1.4).*

The solution flow of (1.4) converges in law to that of the Kunita-type stochastic differential equation written in Itô form as

$$dX_t = W(X_t, dt) + G(X_t)\,dt + \bar{F}_0(X_t)dt, \qquad (1.7)$$

---

[4] But on the other hand it *is* satisfied for systems with superlinear dissipation. This even includes the Allen–Cahn equation on the torus in dimension $d \leq 3$ driven by space-time white noise, as can be deduced from the results in [35–37,56,69]!

where $\bar{F}_0(x) = \int F_0(x, y)\mu(dy)$, $G_i(x) = (\partial_j^{(2)}\Sigma_{ji})(x, x)$, W is a Gaussian random field with correlation

$$\mathbf{E}(W_i(x, t)W_j(\bar{x}, \bar{t})) = (t \wedge \bar{t})\big(\Sigma_{ij}(x, \bar{x}) + \Sigma_{ji}(\bar{x}, x)\big), \tag{1.8}$$

and where $\partial_j^{(2)}$ denotes differentiation in the jth direction of the second argument.

Proof. As already suggested, this is a special case of our main result, Theorem 2.8 below. The fact that Assumptions 2.1–2.3 and 2.7 are implied by Assumption 1.3 is immediate. (Take $E_n = \mathrm{Lip}(\mathcal{Y})$ for every n.)   □

As a consequence, we also have the following functional CLT.

Corollary 1.6. *Let $H \in (\frac{1}{3}, 1)$ and let Assumption 1.3 hold (or let Assumptions 2.1–2.3 hold and when $H > \frac{1}{2}$, let $\int F_i(y)\mu(dy) = 0$ for every $i \geq 1$.)*

Then the stochastic process $Z_t^\varepsilon = \sqrt{\varepsilon}\int_0^{t/\varepsilon} F(Y_r)\,dB_r$ converges to a Wiener process W, weakly in $\mathcal{C}^\alpha([0, T])$ for any $\alpha < \frac{1}{2} \wedge H$. Furthermore, defining the random smooth function $Z_t^{\varepsilon,\delta} = \sqrt{\varepsilon}\int_0^{t/\varepsilon} F(Y_r)\,dB_r^\delta$ with $B^\delta = \phi_\delta * B$, its iterated integral satisfies

$$\lim_{\varepsilon \to 0}\lim_{\delta \to 0}\int_s^t \big(Z_r^{\varepsilon,\delta} - Z_s^{\varepsilon,\delta}\big) \otimes dZ_r^{\varepsilon,\delta} = \int_s^t (W_r - W_s) \otimes dW_r + \Sigma(t - s),$$

where the matrix Σ is given by (1.6) (which is independent of $x, \bar{x}$ in this case).

Remark 1.7. Theorem 1.5 characterises $\lim_{\varepsilon \to 0}\lim_{\delta \to 0} X^{\varepsilon,\delta}$ and shows that it is a Markov process with generator $\mathcal{A}$ given by

$$(\mathcal{A}g)(x) = \sum_{i,j=1}^d \partial_j\big(\Sigma_{ji}(x, \cdot)\partial_i g\big)(x) + \sum_{i=1}^d \bar{F}_0^i(x)\partial_i g(x). \tag{1.9}$$

Our proof actually carries over with minor modifications to the case when $\varepsilon \to 0$ for fixed δ (but with convergence bounds that are uniform in δ!), in which case the limit is given by the same expression (1.7), but with the matrix Σ given by

$$\Sigma_\delta(x, \bar{x}) = \sum_{k=1}^m \int_0^\infty R_\delta(t) \int F_k(x, y) \otimes \big(P_t F_k\big)(\bar{x}, y)\,\mu(dy)\,dt, \tag{1.10}$$

where $R_\delta(t) = \delta^{2H-2}\mathbf{E}v(0)v(t/\delta)$ and $P_t = e^{-\mathcal{L}t}$ denotes the Markov semigroup for Y. We will derive this formula in Sect. 1.3 where we will also see that, for frozen values of x, it is a special case of the Green–Kubo formula [42,45,61]. Note that (1.7)–(1.10) (in particular the convergence of the flow) is also consistent with [22, Theorem 4.3] where a somewhat analogous situation is considered. It follows from Definition 1.2 that $\Sigma_\delta \to \Sigma$ as $\delta \to 0$, so that the two limits commute (in law).

Remark 1.8. There has recently been a surge in interest in the study of slow/fast systems involving fractional Brownian motion. We already mentioned the averaging result [31] which considers the case $H > \frac{1}{2}$ but with $\bar{F} \neq 0$. The work [60] considers the case $H \in (\frac{1}{3}, 1)$ like the present article, but with the very strong assumption that F is independent of the fast variable, in which case only F_0 exhibits rapid fluctuations and one essentially recovers classical averaging results. In [2], the authors consider the case $H > \frac{1}{2}$, but with F independent of the slow variable x and, as in [31], not necessarily averaging to zero. They obtain a description of the fluctuations for (a generalisation of) such systems in the regime where there is an additional small parameter in front of F.

Formula (1.6) holds for the continuum of parameters $H \in (\frac{1}{3}, 1)$. There are two special cases that were previously known. The case $H = \frac{1}{2}$ reduces of course to the classical stochastic averaging results [17,30,66,67] which state that the generator of the limiting diffusion is obtained by averaging the generator for the slow diffusion with the x variable frozen against the invariant measure for the fast process. Note that for this to match (1.9) one needs to interpret the stochastic integral in (1.4) in the Stratonovich sense. This is natural given that this is the interpretation that one obtains when replacing B by a smooth approximation, which is consistent with Remark 1.7. The fact that one also has convergence of flows however (in the case without feedback considered here) appears to be new even in this case.

Another set of closely related classical results deals with "time homogenisation", also known as the Kramers–Smoluchowski limit or diffusion creation [44,61]. There, one considers random ODEs of the type

$$\frac{dX_t^\varepsilon}{dt} = \frac{1}{\sqrt{\varepsilon}} F(X_t^\varepsilon, Y_t^\varepsilon) + F_0(X_t^\varepsilon, Y_t^\varepsilon), \tag{1.11}$$

with F averaging to zero against the stationary measure μ for the fast process Y. In this case, one also obtains a Markov process in the limit $\varepsilon \to 0$ and its generator coincides with (1.9) if one sets $H = 1$. This can be understood by noting that, at least formally, fractional Brownian motion with Hurst parameter $H = 1$ is given by $B(t) = ct$ with c a normal random variable, so that (1.4) reduces to (1.11), except for the random constant c, which then appears quadratically in (1.9) and therefore disappears when averaged out.

The standard proofs of averaging/homogenisation results found in the literature tend to fall roughly into two groups. The first contains functional analytic proofs based on general methods for studying singular limits of the form $\exp(t\mathcal{L}_\varepsilon)$ for $\mathcal{L}_\varepsilon = \varepsilon^{-1}\mathcal{L}_0 + \mathcal{L}_1$. This of course requires the full process (slow plus fast) to be Markovian and completely breaks down in our situation. The second group consists of more probabilistic arguments, which typically rely on using corrector techniques to construct sufficiently many martingales to be able to exploit the well-posedness of the martingale problem for the limiting Markov process. The latter are in principle more promising in our situation since the limiting process is still Markovian, but the lack of Markov property makes it unclear how to construct martingales from our process. (But see Sect. 5.3 for a construction that does go in this direction.)

Instead, our proof relies on rough paths theory [15,53], which has recently been used to recover homogenisation results (formally corresponding to the case $H = 1$), for example in [40]. See also [1,7,12,14] for more recent results with a similar flavour. In the case when the *fast* dynamics is non-Markovian and solves an equation driven by a fractional Brownian motion, a collection of homogenisation results were obtained in [23–26], while stochastic averaging results with non-Markovian fast motions are obtained in [51,52] for the case $H > \frac{1}{2}$. The former group of results are proved using rough path techniques, but there is of course an extensive literature on functional limit theorems based on either central or non-central limit theorems, see for example [3,5,6, 10,54,62,63].

Finally, note that many physical systems can be regarded as a slow/fast systems, this includes second order Langevin equations and tagged particles in a turbulent random field [4,8,16,41,42,61,68]. They also arise in the context of perturbed completely integrable Hamiltonian systems [20,49] and geometric stochastic systems [21,48,50,59]. See also [13,47,70] for some review articles/monographs.

Remark 1.9. It may be surprising that, when $H < \frac{1}{2}$, even though X_ε is driven by a *fractional* Brownian motion and $F(x, y)$ isn't assumed to be centred in the y variable, the limit $\bar{X}$ is a regular diffusion. This is unlike the case $H > \frac{1}{2}$ [31,52] where a non-centred F leads to an averaging result with a process driven by fBm in the limit. This change in behaviour can be understood heuristically as follows. With η as in (3.5), the covariance of $t \mapsto f(Y_t^\varepsilon)\dot{B}_t$ is given by $\zeta_\varepsilon(t - s) = \eta''(t - s)g((t - s)/\varepsilon)$ for some function $g(t) = \langle f, P_t f \rangle$, that would typically converge quite fast to a non-zero limit. The scaling properties of η then show that

$$\int_{\mathbf{R}} \eta''(t)g(t/\varepsilon)\, dt = \varepsilon^{2H-1} \int_{\mathbf{R}} \eta''(t)g(t)\, dt = C_g \varepsilon^{2H-1},$$

for some constant C_g which has no reason to vanish in general. As a consequence, $\varepsilon^{1-2H}\zeta_\varepsilon$ converges pointwise to 0 while its integral remains constant, suggesting that $\varepsilon^{\frac{1}{2}-H} f(Y_t^\varepsilon)\dot{B}_t$ indeed converges to a white noise. When $H > \frac{1}{2}$ however, η'' is not absolutely integrable at infinity and one needs to assume that g vanishes there, which leads to a centering condition. A similar transition from diffusive to super-diffusive behaviour at $H = \frac{1}{2}$ was observed in a different context in [41].

Remark 1.10. As explained, our result implies more, namely that the (random) flow induced by the SDE (1.4) converges in law to that induced by the Kunita-type SDE [46]

$$dx_i = W_i(x, dt) + \left(\partial_j^{(2)}\Sigma_{ji}\right)(x, x)\, dt + \bar{F}_0^i(x)\, dt. \tag{1.12}$$

In other words, the flows $\psi_{s,t}^\varepsilon$, where $\psi_{s,t}^\varepsilon(x)$ denotes the solution at time t to the x-component of (1.4) with initial condition x at time s, converge to a limit $\psi_{s,t}$ which is Markovian in the sense that $\psi_{s,t}$ and $\psi_{u,v}$ are independent whenever $[s, t) \cap [u, v) = \emptyset$. This remark appears to be novel even when $H = \frac{1}{2}$, but it is unclear whether it extends to the case when x feeds back into the dynamic of y as in (1.2).

Remark 1.11. The term $\partial_j^{(2)}\Sigma_{ji}$ appearing in (1.12) looks "almost" like an Itô-Stratonovich correction. In fact, when $\mathcal{L}$ is self-adjoint on $L^2(\mu)$, one has $\Sigma_{ij}(x, \bar{x}) = \Sigma_{ji}(\bar{x}, x)$ in which case (1.12) is equivalent to $dx_i = W_i(x, \circ dt) + \bar{F}_0^i(x)\, dt$.

1.3. Heuristics for general slow/fast random ODEs. We now show how to heuristically derive (1.10). Consider a random ODE of the form

$$\frac{dX_t^\varepsilon}{dt} = \frac{1}{\sqrt{\varepsilon}}\hat{F}(X_t^\varepsilon, Z_t^\varepsilon), \tag{1.13}$$

where $Z_t^\varepsilon = Z(t/\varepsilon)$ for some stationary (but not necessarily Markovian!) stochastic process Z and $\hat{F}(x, \cdot)$ is assumed to be centred with respect to the stationary measure of Z. In the case when $\hat{F}(x, z) = \hat{F}(z)$ does not depend on x, it follows from the Green–Kubo formula [42,45,61] that, at least when Z has sufficiently nice mixing properties, X^ε converges as $\varepsilon \to 0$ to a Wiener process with covariance $\Sigma + \Sigma^\top$, where

$$\Sigma = \int_0^\infty \mathbf{E}\big(\hat{F}(Z_0) \otimes \hat{F}(Z_t)\big)\, dt.$$

This suggests that a natural quantity to consider in the general case is

$$\Sigma(x, \bar{x}) = \int_0^\infty \mathbf{E}\big(\hat{F}(x, Z_0) \otimes \hat{F}(\bar{x}, Z_t)\big) \, dt, \tag{1.14}$$

and that the limit of X^ε as $\varepsilon \to 0$ is a diffusion with generator of the form

$$\big(\mathcal{A}g\big)(x) = \Sigma_{ij}(x, x)\partial_{ij}^2 g(x) + b_i(x)\partial_i g(x), \tag{1.15}$$

for some drift term b.

To derive the correct expression for the drift b, we note that one expects

$$\mathbf{E}\big(X_{t+\delta t}^\varepsilon - X_t^\varepsilon \mid \mathcal{F}_t\big) \approx \delta t \, b(X_t^\varepsilon),$$

in the regime $\varepsilon \ll \delta t \ll 1$. The left-hand side of this expression is given by

$$\frac{1}{\sqrt{\varepsilon}} \int_t^{t+\delta t} \mathbf{E}\big(\hat{F}(X_s^\varepsilon, Z_s^\varepsilon) \mid \mathcal{F}_t\big) \, ds. \tag{1.16}$$

To lowest order, one can approximate this expression by replacing X_s^ε by X_t^ε, but the resulting expression vanishes rapidly for $s \gtrsim t + \varepsilon$ due to the centering condition on $\hat{F}$. To the next order, one has

$$
\begin{aligned}
\mathbf{E}\big(\hat{F}(X_s^\varepsilon, Z_s^\varepsilon) \mid \mathcal{F}_t\big) &\approx \mathbf{E}\Big(\hat{F}\Big(X_t^\varepsilon + \frac{1}{\sqrt{\varepsilon}} \int_t^s \hat{F}(X_t^\varepsilon, Z_r^\varepsilon) \, dr, Z_s^\varepsilon\Big) \,\Big|\, \mathcal{F}_t\Big) \\
&\approx \mathbf{E}\big(\hat{F}(X_t^\varepsilon, Z_s^\varepsilon) \mid \mathcal{F}_t\big) + \frac{1}{\sqrt{\varepsilon}} \int_t^s \mathbf{E}\big(D\hat{F}(X_t^\varepsilon, Z_s^\varepsilon)\hat{F}(X_t^\varepsilon, Z_r^\varepsilon) \mid \mathcal{F}_t\big) \, dr \\
&\approx \sqrt{\varepsilon} \int_0^\infty \mathbf{E}\big(D\hat{F}(X_t^\varepsilon, Z_u)\hat{F}(X_t^\varepsilon, Z_0) \mid \mathcal{F}_t\big) \, du, \tag{1.17}
\end{aligned}
$$

where the last identity follows from the substitution $u = (s - r)/\varepsilon$ combined with the fact that, provided that Z is sufficiently rapidly mixing, we expect the main contribution from this integral to come from $|u| \approx 1$, while typical values of s are such that $(s - t)/\varepsilon \approx \delta t/\varepsilon \gg 1$. Combining this with (1.16) eventually yields the expression

$$b(x) = \int_0^\infty D\hat{F}(x, Z_s)\hat{F}(x, Z_0) \, ds.$$

Comparing this with (1.14), we conclude that

$$b_i(x) = \big(\partial_j \Sigma_{ji}(x, \cdot)\big)(x),$$

(summation over repeated indices is implied) so that (1.15) can be written as

$$\big(\mathcal{A}g\big)(x) = \partial_j\big(\Sigma_{ji}(x, \cdot)\partial_i g\big)(x),$$

which does coincide with the expression (1.9) as desired.

In order to link this calculation with the setting of the previous section, we note that (1.5) (with $F_0 = 0$ for simplicity) can be coerced into the form (1.13) by setting $Z_t = (\delta^{H-1} v(t/\delta), Y_t)$ as well as $\hat{F}(x, (v, y)) = F(x, y)v$. In this case, one has

$$\mathbf{E}\big(\hat{F}(x, Z_0) \otimes \hat{F}(\bar{x}, Z_t)\big) = \delta^{2H-2} R(t/\delta) \sum_{k=1}^{m} \mathbf{E}\big(F_k(x, Y_0) \otimes F_k(\bar{x}, Y_t)\big)$$

$$= R_\delta(t) \sum_{k=1}^{m} \int F_k(x, y) \otimes \big(P_t F_k\big)(\bar{x}, y) \, \mu(dy),$$

so that one does indeed recover the expression (1.10) for any fixed δ.

Remark 1.12. The eagle-eyed reader will have spotted that since the stationary measure of Z is $\mathcal{N}(0, C) \otimes \mu$ for some multiple C of the identity matrix and since $\hat{F}(x, (v, y))$ is linear in v, the centering condition for $\hat{F}$ is *always* satisfied, independently of the choice of F. This explains why our main result does not require any centering condition when $H \leq \frac{1}{2}$. When $H > \frac{1}{2}$ however, the covariance function R decays too slowly for the heuristic derivation just given to apply. The centering condition for F then guarantees that correlations decay sufficiently fast to justify the second step in (1.17).

The remainder of this article is structured as follows. In Sect. 2 we introduce the assumptions on the nonlinearities F_i as well as the fast process Y, we discuss a few examples, and we give provide the statement of our main result. In Sect. 3 we then show that solutions to (1.5) converge as $\delta \to 0$, which yields in particular a precise interpretation of what we mean by (1.4) when $H < \frac{1}{2}$. The strategy of proof is as follows. Given a smooth mollification B^δ of B, we first show convergence of $\int_s^t \int_s^r f(u) \, \dot{B}^\delta(u) du \, g(r) \, \dot{B}^\delta(r) dr$ as $\delta \to 0$ for any deterministic H-Hölder continuous functions f, g. While we are able to reduce this to existing criteria for canonical rough path lifts of Gaussian processes [9,18] in the case where the two fractional Brownian motions appearing in this expression are independent, the case where they are equal requires a bit more care and relies on a simple trick given in Proposition 3.4, which is of independent interest. This then allows us to build an infinite-dimensional rough path $\mathbf{Z}^\varepsilon$ (taking values in a space of vector fields on $\mathbf{R}^d$) associated to (1.4) in a similar way as in [40, Sec. 1.5] (see also the "nonlinear rough paths" of [58] and [26]) and to reformulate (1.4) as an RDE driven by $\mathbf{Z}^\varepsilon$ with nonlinearity given by point evaluation. Section 3.2 provides details of the construction of $\mathbf{Z}^\varepsilon$, while Sect. 3.3 then uses it to formulate our main technical result, namely Theorem 3.14 which shows that $\mathbf{Z}^\varepsilon$ converges to a certain rough path lift of an infinite-dimensional Wiener process with covariance function given by Σ. The remainder of the article is devoted to the proof of this convergence statement. Section 4 shows tightness of the family $\{\mathbf{Z}^\varepsilon\}_{\varepsilon \leq 1}$, while we identify its limit in Sect. 5. In both sections, the cases $H < \frac{1}{2}$ and $H > \frac{1}{2}$ are treated in a completely different way.

The fact that we have convergence of the full infinite-dimensional rough path allows us to conclude that we do not just have convergence of solutions for fixed initial conditions, but of the full solution flow. One point of note is that there are two separate sources of randomness, namely the Markov process Y and the fractional Brownian motion B. Our convergence result is "annealed" in the sense that our convergence in law requires both sources, but a number of intermediate results are "quenched" in the sense that they hold for almost every realisation of Y. It is an open question whether our final convergence result also holds in the quenched sense.

2. Precise Formulation and Results

In this section, we collect the precise assumptions on the functions F_i as well as the Markov process Y.

Convention. We write $A \lesssim B$ as shorthand for $A \leq KB$ with a constant K that will differ from statement to statement.

2.1. Technical assumptions on the fast variable Y. Throughout the article we fix $H \in (\frac{1}{3}, 1)$ as well as a sequence $(E_n)_{n \geq 0}$ of Banach spaces such that $E_n \subset E_{n+1}$ and $E_n \subset L^1(\mathcal{Y}, \mu)$ for every $n \geq 0$, and such that pointwise multiplication is a continuous operation from $E_0 \times E_n$ into E_{n+1} for every $n \geq 0$. We also write simply E instead of E_0 and assume E contains constant functions. See Sect. 2.2 below for two classes of examples showing what type of spaces we have in mind here.

First, we impose that Y has "nice" ergodic properties in the following sense, which in particular implies that μ is its unique invariant measure on $\mathcal{Y}$.

Assumption 2.1. *Let $N = \infty$ for $H > \frac{1}{2}$ and $N = 2$ for $H \in (\frac{1}{3}, \frac{1}{2}]$. For every $n \in [1, N)$, the semigroup P_t extends to a strongly continuous semigroup on E_n and there exist constants C and $c > 0$ (possibly depending on n) such that, for every $f \in E_n$ with $\int_{\mathcal{Y}} f \, d\mu = 0$, one has*

$$\|P_t f\|_{E_n} \leq Ce^{-ct}\|f\|_{E_n}. \tag{2.1}$$

In the low regularity case, we also assume that the process Y has some sample path continuity when composed with a function in E_2.

Assumption 2.2. *For $H \in (\frac{1}{3}, \frac{1}{2})$ there exists $p_\star > \max\{4d, 12/(3H - 1)\}$ such that for every $f \in E_2$*

$$\|f(Y_t) - f(Y_0)\|_{L^{p_\star}} \leq c\|f\|_{E_2}(t^H \wedge 1) \qquad \forall t \geq 0, \tag{2.2}$$

for some constant $c > 0$.

We also need some integrability.

Assumption 2.3. *For $H \geq \frac{1}{2}$, one has $E_n \subset L^2(\mathcal{Y}, \mu)$ for every $n \geq 0$. For $H < \frac{1}{2}$, one has $E_2 \subset L^2(\mathcal{Y}, \mu)$ and $E \subset L^{p_\star}(\mathcal{Y}, \mu)$.*

Remark 2.4. When combining it with the inclusion of the product, Assumption 2.3 implies that $E \subset \bigcap_{p \geq 1} L^p(\mathcal{Y}, \mu)$ for $H \geq \frac{1}{2}$.

Another consequence of these assumptions is as follows.

Remark 2.5. As a consequence of Assumption 2.2, we conclude that if $f \in E_2$ and $H < \frac{1}{2}$, then

$$\|P_t f - f\|_\mu^2 = \mathbf{E}\big|\mathbf{E}\big(f(Y_t) - f(Y_0) \,|\, Y_0\big)\big|^2 \leq \mathbf{E}|f(Y_t) - f(Y_0)|^2$$
$$\leq \|f\|_{E_2}^2\big(t^{2H} \wedge 1\big).$$

Recalling the definition of $\mathcal{L}^\alpha$ from Definition 1.2, it follows that $E_2 \subset \mathrm{Dom}(\mathcal{L}^\alpha)$ for every $\alpha < H$ so that (1.6) is indeed well defined provided that $F_k(x, \cdot) \in E_2$ for every x. This will be guaranteed by Assumption 2.7 below.

2.2. Examples of fast variables . One possible concrete framework is as follows. Fix two weights $V : \mathcal{Y} \to [1, \infty]$ and $W : \mathcal{Y} \to (0, \infty)$ and a metric d on $\mathcal{Y}$ generating its topology with the property that there exists $C > 0$ such that, for all $x, y \in \mathcal{Y}$ with $d(x, y) \le 1$, one has

$$V(x) \le CV(y), \qquad W(x) \le CW(y). \tag{2.3}$$

For $n \ge 1$, we then let $\mathcal{B}_{V,W}$ be the Banach space of functions $f : \mathcal{Y} \to \mathbf{R}$ such that

$$\|f\|_{V,W} \overset{\text{def}}{=} \sup_{x \in \mathcal{Y}} \frac{|f(x)|}{V(x)} + \sup_{\substack{x, y \in \mathcal{Y} \\ d(x,y) \le 1}} \frac{|f(x) - f(y)|}{d(x, y)W(x)V(x)} < \infty.$$

One choice of scale of function spaces that is suitable for a large class of Markov processes is to take $E_n = \mathcal{B}_{V,W}$ for every $n \ge 1$ (and suitably chosen V and W), while E_0 is chosen be the space of bounded Lipschitz continuous functions, namely $\mathcal{B}_{1,1}$.

This framework is relatively general since it allows for a wide variety of choices of V, W, and of distance functions on $\mathcal{Y}$, see [33,36]. For example, it was shown in [32, Thm. 1.4] that the 2D stochastic Navier–Stokes equations exhibit a spectral gap in such spaces under extremely weak conditions on the driving noise. More precisely, for every η small enough there exist constants C and γ such that

$$\left\| P_t f - \int f \, d\mu \right\|_\eta \le Ce^{-\gamma t} \|f\|_\eta,$$

for every Fréchet differentiable function f for every $t \ge 0$, where

$$\|f\|_\eta = \sup_x e^{-\eta|x|^2} \big(|f(x)| + |Df(x)| \big). \tag{2.4}$$

This at first sight appears to fall outside our framework, but one notices that if one sets

$$d(x, y) = \inf_{\gamma : x \to y} \int_0^1 (1 + |\gamma(t)|) |\dot{\gamma}(t)| \, dt, \tag{2.5}$$

then the norm $\| \cdot \|_{V,W}$ with $V(x) = \exp(\eta|x|^2)$ and $W(x) = 1/(1 + |x|)$ is equivalent to the norm (2.4). The reason for the choice of d as in (2.5), which is then "undone" by our choice of W, is to guarantee that (2.3) holds for V, which would not be the case for $|x - y| \le 1$ in the Euclidean distance.

To verify Assumption 2.2 one can then for example make use of the following.

Lemma 2.6. *Suppose that* $\int \big(V(x)(1 + W(x)) \big)^{p_\star} \mu(dx) < \infty$ *and there exists a constant* c *such that, for some* $\alpha_0 > 1 - 2H$,

$$\|d(Y_t, Y_0)\|_{L^{p_\star}} \le c \big(t^{\alpha_0} \wedge 1 \big) \qquad \forall t \ge 0. \tag{2.6}$$

Let $f \in \mathcal{B}_{V,W}$ *and* $2p \le p_\star$, *then*

$$\|f(Y_t) - f(Y_0)\|_{L^p} \lesssim \|f\|_{V,W} (1 \wedge t^{\alpha_0}).$$

In particular, on any fixed time interval, we have $\mathbf{E}\|f(Y_\cdot)\|_{\mathcal{C}^\alpha}^p < \infty$ *provided that* $p(\alpha_0 - \alpha) > 1$.

Proof. For $f \in E$ with $\|f\|_E \leq 1$ and for $p \leq \frac{1}{2} p_\star$, one has

$$
\begin{aligned}
\|f(Y_t) - f(Y_0)\|_{L^p} &\leq \|W(Y_0)V(Y_0)d(Y_t, Y_0) + \mathbf{1}_{d(Y_t, Y_0) > 1}\|f\|_E\big(V(Y_0) + V(Y_t)\big)\|_{L^p} \\
&\lesssim \|(1 + W(Y_0))V(Y_0)\|_{L^{2p}}\Big((1 \wedge t^\alpha) + \mathbf{P}(d(Y_t, Y_0) > 1)^{\frac{1}{2p}}\Big) \\
&\lesssim 1 \wedge t^\alpha,
\end{aligned}
$$

where we combined (2.6) with Markov's inequality in the last step. $\qquad\square$

When $H \geq \frac{1}{2}$, Assumption 2.2 is empty, so only integrability conditions are required on the spaces E_n. This allows for example to use Harris's theorem [29, 34, 55] to verify Assumption 2.1 for spaces of functions with weighted supremum norms. More precisely, one would then take E to be the space of all bounded Borel measurable functions and $E_n = \mathcal{B}_V$, the Banach space of functions $f : \mathcal{Y} \to \mathbf{R}$ such that

$$
\|f\|_V \overset{\text{def}}{=} \sup_{x \in \mathcal{Y}} \frac{|f(x)|}{V(x)} < \infty.
$$

In order to verify our assumptions, it then suffices that V is a square integrable Lyapunov function for the Markov process Y and that the sublevel sets of V satisfy a 'small set' condition for the transition probabilities of Y [55].

2.3. Main results. One final assumption we need is that the nonlinearities F and F_0 appearing in (1.4) are sufficiently nice E-valued functions of their first argument. More precisely, we assume the following.

Assumption 2.7. *The map $x \mapsto F(x, \cdot)$ is of class $\mathcal{C}^4$ with values in E and there exists an exponent $\kappa > \frac{16d}{p_\star}$ with $p_\star$ as in Assumption 2.2 (and simply $\kappa > 0$ when $H > \frac{1}{2}$) such that, for every multi-index ℓ of length at most 4,*

$$
\|D_x^\ell F(x, \cdot)\|_E \lesssim (1 + |x|)^{-\kappa}.
$$

The same is assumed to hold true for F_0. When $H > \frac{1}{2}$, we further assume that $\int F_i(x, y)\,\mu(dy) = 0$ for every $i \neq 0$ and every x.

The condition $F \in \mathcal{C}^4$ is of course suboptimal and could probably be lowered to $F \in \mathcal{C}^\beta$ for $\beta > \max\{H^{-1}, 2\}$ and $F_0 \in \mathcal{C}^\beta$ for $\beta > 1$, at least if enough integrability is assumed in Assumption 2.3. We also now fix a Schwartz function ϕ integrating to 1 and set $\phi_\delta(t) = \frac{1}{\delta}\phi(t/\delta)$. We then write B^δ for the convolution of B with this mollifier, namely

$$
B^\delta(t) = \frac{1}{\delta}(\phi_\delta * B)(t) = \frac{1}{\delta}\int_{\mathbf{R}} \phi\Big(\frac{t - s}{\delta}\Big)B(s)\,ds.
$$

With this notation, the solutions to (1.5) are equal in law to the process given by

$$
\dot{X}_t^{\varepsilon, \delta} = \varepsilon^{\frac{1}{2} - H}F(X_t^{\varepsilon, \delta}, Y_t^\varepsilon)\,\dot{B}^{\varepsilon\delta} + F_0(X_t^{\varepsilon, \delta}, Y_t^\varepsilon). \tag{2.7}
$$

Since $B^{\varepsilon\delta}$ is smooth, this equation should be interpreted as an ordinary differential equation that just happens to have random coefficients. With all these preliminaries at hand, our main result is the following.

Theorem 2.8. *For $H \in (\frac{1}{3}, 1]$ and under Assumptions 2.1–2.3, and 2.7, the conclusions of Theorem 1.5 hold. With $X^{\varepsilon,\delta}$ defined in (2.7), the convergence $X^{\varepsilon,\delta} \to X^{\varepsilon}$ furthermore holds in probability.*

Proof. The convergence in probability of the flow $X^{\varepsilon,\delta} \to X^{\varepsilon}$ is the content of Proposition 3.2 below. The proof of the conclusion of Theorem 1.5, namely the convergence in law of the flow for (1.4) as $\varepsilon \to 0$ is the content of Corollary 3.15. □

3. Convergence of Smooth Approximations

We first address the question of the convergence in probability of solutions to (1.5) to those of (1.4) as $\delta \to 0$ for $\varepsilon > 0$ fixed. In fact, we will directly provide an interpretation of (1.4) and show that this interpretation is sufficiently stable to allow for the approximation (1.5).

Our convergence proof relies on the theory of rough paths; we refer to [15] for an introduction. The main insight of this theory is that even though, for $H \leq \frac{1}{2}$, the solution map $B \mapsto X$ for equations of the type (1.4) isn't continuous when viewing B as an element of any classical function space large enough to contain typical sample paths of fractional Brownian motion, it does become continuous when enhancing B with its iterated integrals $\mathbb{B} = \int B \otimes dB$ and endowing the space of pairs $(B, \mathbb{B})$ with a suitable topology.

For this, consider for any $x \in \mathbf{R}^d$ the processes

$$Z^{\varepsilon,\delta}_{s,t}(x) = \varepsilon^{\frac{1}{2}-H} \int_s^t F(x, Y^{\varepsilon}_r) \, dB^{\delta}(r), \quad \bar{Z}^{\varepsilon}_{s,t}(x) = \int_s^t F_0(x, Y^{\varepsilon}_r) \, dr. \qquad (3.1)$$

Here, the first integral is interpreted as a Wiener integral which makes sense also when $\delta = 0$ and, when $\delta > 0$, coincides with the Riemann–Stieltjes integral. Recall that the Wiener integral of a deterministic (or independent) integrand against any Gaussian process B is well-defined provided that the integrand belongs to the reproducing kernel Hilbert space $\mathcal{H}_B$ of B and provides an isometric embedding $\mathcal{H}_B \ni f \mapsto \int f \, dB \in L^2(\Omega, \mathbf{P})$. In the case of fractional Brownian motion, it is known that $L^2 \subset \mathcal{H}_B$ when $H \geq \frac{1}{2}$ while for $H < \frac{1}{2}$ one has $\mathcal{C}^{\alpha} \subset \mathcal{H}_B$ for every $\alpha > \frac{1}{2} - H$. The fact that for fixed x and $\varepsilon > 0$, $t \mapsto F(x, Y^{\varepsilon}_t)$ belongs to $\mathcal{H}_B$ for all $H > \frac{1}{3}$ is then a simple consequence of Assumptions 2.2 and 2.3 combined with Kolmogorov's continuity criterion (when $H < \frac{1}{2}$).

Write $\mathcal{B} = \mathcal{C}^3_b(\mathbf{R}^d, \mathbf{R}^d)$ and $\mathcal{B}_k = \mathcal{C}^3_b(\mathbf{R}^{d \cdot k}, (\mathbf{R}^d)^{\otimes k})$, so that one has canonical inclusions of the algebraic tensor product $\mathcal{B}_k \otimes_0 \mathcal{B}_\ell \subset \mathcal{B}_{k+\ell}$ with the usual identification $(f \otimes g)(x, y) = f(x)g(y)$ thanks to the fact that $\|f \otimes g\|_{\mathcal{B}_{k+\ell}} \leq \|f\|_{\mathcal{B}_k} \|g\|_{\mathcal{B}_k}$. Given a final time $T > 0$ and $\alpha \in (\frac{1}{3}, \frac{1}{2})$, we define the space $\mathcal{C}^{\alpha}([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ of α-Hölder rough paths in the usual way [15, Def. 2.1], but with all norms of level-2 objects in $\mathcal{B}_2$. Recall that an α-Hölder rough path $(X, \mathbb{X})$ is a pair of functions where $X \in \mathcal{C}^{\alpha}([0, T], \mathcal{B})$ with $X_0 = 0$ and $\mathbb{X} : \Delta_T \to \mathcal{B}_2$ where $\Delta_T := \{(s, t) : 0 \leq s \leq t \leq T\}$ is the two-simplex with $|\mathbb{X}|_{2\alpha} := \sup_{(s,t) \in \Delta_T} \frac{\|\mathbb{X}_{s,t}\|_{\mathcal{B}_2}}{|t-s|^{2\alpha}} < \infty$. In addition, Chen's relation is imposed, namely $\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t}$.

We define the second-order processes $\mathbb{Z}^{\varepsilon,\delta}$ and $\bar{\mathbb{Z}}^{\varepsilon}$ by

$$\mathbb{Z}^{\varepsilon,\delta}_{s,t}(x, \bar{x}) = \int_s^t Z^{\varepsilon,\delta}_{s,r}(x) \, dZ^{\varepsilon,\delta}_{s,r}(\bar{x}), \qquad \bar{\mathbb{Z}}^{\varepsilon}_{s,t}(x, \bar{x}) = \int_s^t \bar{Z}^{\varepsilon}_{s,r}(x) \, d\bar{Z}^{\varepsilon}_{s,r}(\bar{x}),$$

(the differentials are taken in the r variable) and we define $\mathbf{Z}^{\varepsilon,\delta} = (Z^{\varepsilon,\delta}, \mathbb{Z}^{\varepsilon,\delta})$, $\bar{\mathbf{Z}}^{\varepsilon} = (\bar{Z}^{\varepsilon}, \bar{\mathbb{Z}}^{\varepsilon})$. Note here that $r \mapsto Z^{\varepsilon,\delta}_{s,r}(x)$ is smooth and $r \mapsto \bar{Z}^{\varepsilon}_{s,r}(x)$ is Hölder continuous for any exponent strictly less than 1, so these integrals should be interpreted as regular Riemann–Stieltjes integrals. In Sect. 3.2 below we will give a proof of the following result.

Proposition 3.1. *Let* $H \in (\frac{1}{3}, 1]$, *let* $\alpha \in (\frac{1}{3}, H \wedge \frac{1}{2})$ *and* $\beta \in (1 - \alpha, 1)$, *and let Assumptions* 2.1–2.3, *and* 2.7 *hold. Then,* $\mathbf{Z}^{\varepsilon,\delta}$ *and* $\bar{\mathbf{Z}}^{\varepsilon}$ *admit versions that are random elements in* $\mathscr{C}^{\alpha}([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ *and* $\mathscr{C}^{\beta}([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ *respectively. Furthermore,* $\mathbf{Z}^{\varepsilon,\delta}$ *converges in probability in* $\mathscr{C}^{\alpha}([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ *as* $\delta \to 0$ *to the random rough path* $\mathbf{Z}^{\varepsilon}$ *characterised in Proposition* 3.11 *below. (In particular, the first order component* Z^{ε} *of* $\mathbf{Z}^{\varepsilon}$ *is given, for any fixed* x, *by the Wiener integral* (3.1) *with* $\delta = 0$.)

For now, we take this result as granted. With this result in place, we obtain the following convergence result as $\delta \to 0$.

Proposition 3.2. *The second claim of Theorem* 2.8 *holds.*

Proof. With the space $\mathcal{B}$ as above, let $\delta \colon \mathbf{R}^d \to L(\mathcal{B}, \mathbf{R}^d)$ be the function given by $\delta(x)(f) = f(x)$. We then claim that, for any $\varepsilon, \delta > 0$, (2.7) can be rewritten as the rough differential equation (RDE) driven by the infinite-dimensional rough paths $\mathbf{Z}^{\varepsilon,\delta}$ and $\bar{\mathbf{Z}}^{\varepsilon}$ defined above and given by

$$dX = \delta(X) \, d\mathbf{Z}^{\varepsilon,\delta} + \delta(X) \, d\bar{\mathbf{Z}}^{\varepsilon}. \tag{3.2}$$

Note that since $\alpha + \beta > 1$, there is no need to specify cross-integrals between $\mathbf{Z}^{\varepsilon,\delta}$ and $\bar{\mathbf{Z}}^{\varepsilon}$ since they can be defined in a canonical way using Young integration [72].

To check that this RDE is well-posed for any rough paths $\mathbf{Z}^{\varepsilon,\delta}$ and $\bar{\mathbf{Z}}^{\varepsilon}$ belonging to $\mathscr{C}^{\alpha}([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ and $\mathscr{C}^{\beta}([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ respectively, we note first that one readily verifies that the map δ is Fréchet differentiable, and actually even $\mathcal{C}^3_b$. Its differential $D\delta$ at $x \in \mathbf{R}^d$ in the direction of $y \in \mathbf{R}^d$ is given by $(D\delta)_x(y)(h) = (Dh)_x(y)$ where $h \in L(\mathcal{B}, \mathbf{R}^d)$. Morerover $|(D\delta)_x(h)| \leq |h|_{\mathcal{B}}$. In particular we may consider the map $D\delta \cdot \delta \colon \mathbf{R}^d \to L(\mathcal{B} \otimes \mathcal{B}, \mathbf{R}^d)$ which for $h = h_1 \otimes h_2 \in \mathcal{B}_2$ is given by

$$\begin{aligned}
\big(D\delta \cdot \delta\big)(x)h &= (D\delta)_x\big(\delta(x)(h_1)\big)(h_2) = (D\delta)_x\big(h_1(x)\big)(h_2) \\
&= (Dh_2)_x(h_1(x)) = (\operatorname{tr} D^{(2)}h)(x, x), \tag{3.3}
\end{aligned}$$

for a suitable partial trace tr. This shows that $D\delta \cdot \delta$ extends continuously to a $\mathcal{C}^2_b$ map from $\mathbf{R}^d$ into $L(\mathcal{B}_2, \mathbf{R}^d)$. (Since $\mathcal{B}_2$ differs from the projective tensor product of $\mathcal{B}$ with itself, this doesn't automatically follow from the fact that δ itself is $\mathcal{C}^3_b$.) Retracing the standard existence and uniqueness proof for RDEs, [15, Sec. 8.5] then shows that (3.2) admits unique (global) solutions for every initial condition and every driving path. Furthermore, if the sample path $t \mapsto Y(t)$ is given by any continuous $\mathcal{Y}$-valued function then, under the stated regularity conditions on F, F_0, it is straightforward to verify that the solutions to (3.2) coincide with those of (2.7).

Since the RDE solution is a jointly locally Lipschitz continuous function of both the initial data x_0 and the driving paths $\mathbf{Z}^{\varepsilon,\delta}$ and $\bar{\mathbf{Z}}^{\varepsilon}$ into $\mathcal{C}^{\alpha}(\mathbf{R}_+, \mathbf{R}^d)$, the claim that $X^{\varepsilon,\delta} \to X^{\varepsilon}$ in probability then follows immediately from Proposition 3.1. $\qquad\square$

Remark 3.3. This shows that it is consistent to *define* solutions to (1.4) in the general case as simply being a shorthand for solutions to (3.2) driven by the pair of rough paths $\mathbf{Z}^{\varepsilon}$ and $\bar{\mathbf{Z}}^{\varepsilon}$. This is the interpretation that we will use from now on. The fact that in the special case $H = \frac{1}{2}$ this coincides with the Stratonovich interpretation of the equation follows as in [15, Thm 9.1].

3.1. Preliminary results. In this section we present a few general results that will be used in the proof of Proposition 3.1. We start with the following elementary property of the second Wiener chaos.

Proposition 3.4. *Let $\mathcal{H} \subset \mathcal{E}$ be an abstract Wiener space and let $B, \tilde{B}$ be two i.i.d. Gaussian random variables on $\mathcal{E}$ with Cameron–Martin space $\mathcal{H}$. Let $K^\delta \colon \mathcal{E} \times \mathcal{E} \to \mathbf{R}$ be continuous bilinear maps such that the limit $\tilde{K} = \lim_{\delta \to 0} K^\delta(B, \tilde{B})$ exists in L^2. Then, the limit*

$$K = \lim_{\delta \to 0} \left(K^\delta(B, B) - \mathbf{E}K^\delta(B, B) \right) \tag{3.4}$$

exists in L^2. Furthermore, the limit in (3.4) depends only on the limit $\tilde{K}$ and not on the approximating sequence K^δ and one has the bound $\mathbf{E}K^2 \leq 2\mathbf{E}\tilde{K}^2$.

Proof. The Gaussian probability space generated by the pair $(B, \tilde{B})$ has Cameron–Martin space $\mathcal{H} \oplus \tilde{\mathcal{H}}$ where $\tilde{\mathcal{H}}$ is a copy of $\mathcal{H}$. Since K^δ is bilinear and $K^\delta(B, \tilde{B})$ has vanishing expectation, it belongs to the second homogeneous Wiener chaos, so that there exists $\hat{K}^\delta \in (\mathcal{H} \oplus \tilde{\mathcal{H}}) \otimes_s (\mathcal{H} \oplus \tilde{\mathcal{H}})$ with $K^\delta(B, \tilde{B}) = \mathcal{I}_2(\hat{K}^\delta)$, where $\mathcal{I}_k$ denotes the usual isometry between kth symmetric tensor power and kth homogeneous Wiener chaos, see [57].

Note now that, interpreting $\hat{K}^\delta$ as a Hilbert–Schmidt operator on $\mathcal{H} \oplus \tilde{\mathcal{H}}$, there exists $\mathrm{K}^\delta \in \tilde{\mathcal{H}} \otimes \mathcal{H}$ such that $\hat{K}^\delta = \iota\mathrm{K}^\delta$, where $\iota \colon \tilde{\mathcal{H}} \otimes \mathcal{H} \to (\mathcal{H} \oplus \tilde{\mathcal{H}}) \otimes_s (\mathcal{H} \oplus \tilde{\mathcal{H}})$ is given by

$$\iota\mathrm{K} = \frac{1}{2} \begin{pmatrix} 0 & \tau\mathrm{K} \\ \mathrm{K} & 0 \end{pmatrix},$$

with the obvious matrix notation and $\tau \colon \tilde{\mathcal{H}} \otimes \mathcal{H} \to \mathcal{H} \otimes \tilde{\mathcal{H}}$ the transposition operator.

This is because the first diagonal block is obtained by testing against $\langle k, (B, \tilde{B}) \rangle$ where $k = (f, 0) \otimes_s (g, 0)$ with $f, g \in \mathcal{H}$, yielding

$$\mathbf{E}\left(K^\delta(B, \tilde{B})\left(\langle f, B \rangle \langle g, B \rangle - \langle f, g \rangle \right) \right) = 0.$$

The second diagonal block vanishes for the same reason with the roles of B and $\tilde{B}$ exchanged. (Here we denote by $x \mapsto \langle x, f \rangle$ the unique element of $L^2(\mathcal{E})$ which is linear on a set of full measure containing $\mathcal{H}$ and coincides with $\langle f, \cdot \rangle$ there. In fact, $\langle x, f \rangle = f^*(x)$ when $f^* \in \mathcal{E}^*$ and $\langle B, f \rangle = f^* \circ B$.) On the other hand, one has

$$K^\delta_\diamond(B, B) \overset{\text{def}}{=} K^\delta(B, B) - \mathbf{E}K^\delta(B, B) = \frac{1}{2}\mathcal{I}_2(\mathrm{K}^\delta + \tau\mathrm{K}^\delta),$$

where this time $\mathcal{I}_2$ refers to the isometry between $\mathcal{H} \otimes_s \mathcal{H}$ and the second chaos generated by B only. Since $\mathcal{I}_2$ and τ are both isometries, it immediately follows that $\mathbf{E}(K^\delta_\diamond(B, B))^2 \leq \|\mathrm{K}^\delta\|^2 = 2\|\iota\mathrm{K}^\delta\|^2 = 2\mathbf{E}(K^\delta(B, \tilde{B}))^2$, and similarly for differences $K^\delta - K^{\delta'}$, so that the claim follows. □

Remark 3.5. It follows immediately from (3.4) that if we replace K^δ by the same sequence of bilinear maps, but with their two arguments exchanged, the limit one obtains is the same.

Before we turn to the precise statement, we introduce the following notation which will be used repeatedly in the sequel. We write η for the distribution on $\mathbf{R}$ given by

$$\langle \eta, \phi \rangle = \int_{\mathbf{R}} |t|^{2H} \phi(t) \, dt, \tag{3.5}$$

and η'' for its second distributional derivative. For $a < 0 < b$, we will then make the abuse of notation $\int_a^b \eta''(t) \phi(t) \, dt$ as a shorthand for $\lim_{\varepsilon \to 0} \langle \eta'', \phi \mathbf{1}_{[a,b]}^{\varepsilon} \rangle$, where $\mathbf{1}_{[a,b]}^{\varepsilon}$ denotes a mollification of the indicator function $\mathbf{1}_{[a,b]}$. The following is elementary.

Lemma 3.6. *Let $a < 0 < b$ and $H \in (\frac{1}{3}, \frac{1}{2})$. Setting $\alpha_H = H(1 - 2H)$ and $\phi_0 = \phi(0)$, the limit above is given by*

$$\int_a^b \eta''(t) \phi(t) \, dt = -2\alpha_H \int_a^b |t|^{2H-2} \left(\phi(t) - \phi_0 \right) dt + 2H\phi_0 \left(|a|^{2H-1} + |b|^{2H-1} \right), \tag{3.6}$$

independently of the choice of mollifier, thus justifying the notation. For $a = 0$, we set

$$\int_0^b \eta''(t) \phi(t) \, dt = -2\alpha_H \int_0^b |t|^{2H-2} \left(\phi(t) - \phi_0 \right) dt + 2H\phi_0 |a|^{2H-1},$$

which can be justified in a similar way, provided that the mollifier one uses is symmetric. (This in turn is the case if we view η as the limit of covariances of smooth approximations to fractional Brownian motion.)

For $H = \frac{1}{2}$, one similarly has $\int_a^b \eta''(t) \phi(t) \, dt = \phi_0$ and $\int_0^b \eta''(t) \phi(t) \, dt = \frac{1}{2}\phi_0$, while for $H \in (\frac{1}{2}, 1)$, η'' is given by the locally integrable function $t \mapsto -2\alpha_H |t|^{2H-2}$. □

We now show that for any fixed $\varepsilon > 0$, the processes Z^{ε} satisfy a suitable form of Hölder regularity. To keep notations shorter, we define the collection of processes

$$Z_{s,t}^f \stackrel{\text{def}}{=} \int_s^t f(r) \, dB_i(r) = \sum_{i=1}^m \int_s^t \langle f(r), e^i \rangle_{\mathbf{R}^m} dB_r^i, \tag{3.7}$$

indexed by $\mathbf{R}^m$-valued functions f that belong to the reproducing kernel space of the fractional Brownian motion B. Here $\{e^i\}$ is an o.n.b. of $\mathbf{R}^m$ and $B_r = (B_r^1, \ldots, B_r^m)$. We start our analysis with some preliminary result for the irregular case $H < \frac{1}{2}$.

Lemma 3.7. *Let $H \in (\frac{1}{3}, \frac{1}{2})$ and let $f \in \mathcal{C}^{\beta}([0, 1], \mathbf{R}^m)$ for some $\beta > (1 - 2H) \vee 0$. The processes Z^f satisfy the Coutin–Qian conditions [9, 18, Def. 14] in the sense that*

$$\mathbf{E}\left(Z_{s,t}^f(x)^2 \right) \lesssim \|f\|_{\infty} \|f\|_{\mathcal{C}^{\beta}} |t - s|^{2H},$$

$$\left| \mathbf{E}\left(Z_{s,s+h}^f(x) \, Z_{t,t+h}^f(x) \right) \right| \lesssim \|f\|_{\infty}^2 |t - s|^{2H-2} h^2,$$

for all $0 < s < t < 1$ and all $h \in (0, t - s]$.

Proof. The mixed second order distributional derivative of $\mathbf{E}(B_s^\delta B_t^\delta)$ is given by $\frac{1}{2}\eta_\delta''(s-t)$, the convolution of $\frac{1}{2}\eta''$ with a symmetric mollifier at scale δ. Mollifying B and taking limits shows that we have the identity

$$\mathbf{E}(Z_{s,t}^f)^2 = \lim_{\delta \to 0} \int_s^t \int_s^t \mathbf{E}(\dot{B}_u^\delta \dot{B}_v^\delta) f(u) f(v) \, du \, dv$$

$$= \frac{1}{2} \int_s^t \int_s^t \eta''(x-y) f(x) f(y) \, dx \, dy,$$

(with summation over the components of f implied). For $H < \frac{1}{2}$, this yields the bound

$$\mathbf{E}(Z_{s,t}^f)^2 = \frac{1}{4} \int_{s-t}^{t-s} \eta''(v) \int_{2s+|v|}^{2t-|v|} \left(f\left(\frac{u+v}{2}\right) f\left(\frac{u-v}{2}\right) \right) du \, dv$$

$$= -\frac{1}{2}\alpha_H \int_{s-t}^{t-s} |v|^{2H-2} \int_{2s+|v|}^{2t-|v|} \left(f\left(\frac{u+v}{2}\right) f\left(\frac{u-v}{2}\right) - f\left(\frac{u}{2}\right)^2 \right) du \, dv$$

$$+ \frac{\alpha_H}{1-2H} \int_0^{t-s} u^{2H-1} \left(f\left(t-\frac{u}{2}\right)^2 + f\left(s+\frac{u}{2}\right)^2 \right) du$$

$$\lesssim \|f\|_{\mathcal{C}^\beta} \|f\|_\infty |t-s|^{2H+\beta} + \|f\|_\infty^2 |t-s|^{2H}, \tag{3.8}$$

as required.

Regarding the covariance, we have

$$\left| \mathbf{E}\left(Z_{s,s+h}^f Z_{t,t+h}^f \right) \right| \lesssim \|f\|_\infty^2 \int_0^h \int_0^h |t-s+v-u|^{2H-2} \, du \, dv$$

$$\lesssim \|f\|_\infty^2 h^2 |t-s|^{2H-2},$$

when $0 < h \le t-s$ so the intervals overlap only at most one point, as required. $\square$

If $\tilde{B}$ is an independent copy of B, we can combine this result with those of [18] to conclude that there is a canonical rough path associated with the path $\left(\int f(r) \, dB(r), \int g(r) \, d\tilde{B}(r) \right)$ for any $f, g \in \mathcal{C}^\beta([0,1], \mathbf{R}^m)$. We now show that there also exists a canonical lift for (Z^f, Z^g), where the integrals are defined with respect to the same fractional Brownian motion B. With this notation, we then have the following result.

Proposition 3.8. *For $H \in (\frac{1}{3}, \frac{1}{2})$, we set $\alpha = H$ and $U = \mathcal{C}^\beta([0,1], \mathbf{R}^m)$ for some $\beta \in (\frac{1}{3}, H)$. For $H \in [\frac{1}{2}, 1)$, we fix $p > 3/(3H-1)$ and set $\alpha = H - \frac{1}{p}$ and $U = L^p([0,1], \mathbf{R}^m)$.*

Then, for any finite collection $\{f_i\}_{i \le N} \subset U$, there is a "canonical lift" of

$$Z^f = (Z^{f_1}, \ldots, Z^{f_N})$$

given by (3.7) to a geometric rough path $\mathbf{Z}^f = (Z^f, \mathbb{Z}^f)$ on $\mathbf{R}^N$ such that

$$\sup_{s \ne t} |t-s|^{-q\alpha} \mathbf{E}\left(|Z_{s,t}^f|^q + |\mathbb{Z}_{s,t}^f|^{\frac{q}{2}} \right) \lesssim \sum_i \|f_i\|_U^q, \tag{3.9}$$

for every $q \ge 1$. This is obtained by taking the limit as $\delta \to 0$ of the canonical lift of the smooth paths $Z^{\delta,f}$ defined as in (3.7) but with B replaced by B^δ.

Remark 3.9. As usual, "geometric" here means that $\mathbf{Z}^f$ is the limit of canonical lifts of smooth functions. Indeed, for B^δ the convolution of B with a mollifier at scale $\delta > 0$, $\mathbb{Z}^f$ is given by

$$(\mathbb{Z}^f_{s,t})_{ij} = \lim_{\delta \to 0} \int_s^t \int_s^r f_i(u)\, \dot{B}^\delta(u)du\, f_j(r)\, \dot{B}^\delta(r)dr,$$

and this limit is independent of the choice of mollifier (and therefore "canonical").

Proof. We only need to show (3.9) for $q = 2$ since Z and $\mathbb{Z}$ belong to a Wiener chaos of fixed order. (Recall that the f_i are considered deterministic here.)

We start with the case $H \in (\frac{1}{3}, \frac{1}{2})$. Let $\tilde{B}$ denote an independent copy of the fractional Brownian motion B and let $\tilde{Z}^{f,g} = (Z^f, \tilde{Z}^g)$ where $\tilde{Z}^g$ is defined like Z^g but with B replaced by $\tilde{B}$. By Lemma 3.7, for any $f, g \in \mathcal{C}^\beta$, we can then apply [18, Thm 35] to construct a second-order process $\tilde{\mathbb{Z}}^{f,g}_{s,t}$ satisfying the Chen identity

$$\tilde{\mathbb{Z}}^{f,g}_{s,t} - \tilde{\mathbb{Z}}^{f,g}_{s,u} - \tilde{\mathbb{Z}}^{f,g}_{u,t} = Z^f_{s,u}\, \tilde{Z}^g_{u,t}.$$

It furthermore coincides with the Wiener integral

$$\tilde{\mathbb{Z}}^{f,g}_{s,t} = \int_s^t Z^f_{s,r}\, d\tilde{Z}^g_{s,r} = \int_s^t \int_s^r f(u)\, dB(u)\, g(r)\, d\tilde{B}(r), \tag{3.10}$$

which makes sense since the Coutin–Qian condition guarantees that $r \mapsto Z^f_{s,r}$ belongs to the reproducing kernel space of Z^g. It is furthermore such that smooth approximations to (3.10) (replace B and $\tilde{B}$ by B^δ and $\tilde{B}^\delta$, obtained by convolution with a mollifier at scale $\delta \to 0$) converge to it in L^2. In particular, $\tilde{\mathbb{Z}}^{f,g}$ belongs to the second Wiener chaos generated by $(B, \tilde{B})$ and is of the form of the limits considered in Proposition 3.4.

We now want to replace $\tilde{B}$ by B. For an approximation B^δ as mentioned above, setting

$$\mathbb{Z}^{\delta,f,g}_{s,t} = \int_s^t \int_s^r f(u)\, dB^\delta(u)\, g(r)\, dB^\delta(r), \tag{3.11}$$

we have

$$\mathbf{E}\mathbb{Z}^{\delta,f,g}_{s,t} = \frac{1}{2} \int_s^t \int_s^r f(u)g(r)\, \eta''_\delta(u-r)\, du\, dr, \tag{3.12}$$

where η_δ is an even δ-mollification of $t \mapsto |t|^{2H}$, so that in particular $\int_0^\infty \eta''_\delta(s)\, ds = 0$, since $\eta'_\delta(0) = 0$. Similarly to (3.8), we can then rewrite (3.12) as

$$\mathbf{E}\mathbb{Z}^{\delta,f,g}_{s,t} = \frac{1}{2} \int_s^t g(r) \int_s^r \big(f(u) - f(r)\big)\eta''_\delta(r-u)\, du\, dr$$
$$- \frac{1}{2} \int_s^t g(r) \int_{-\infty}^s f(r)\eta''_\delta(r-u)\, du\, dr.$$

It follows immediately that one has

$$\lim_{\delta \to 0} \mathbf{E} \mathbb{Z}_{s,t}^{\delta,f,g} = H(2H-1) \int_s^t g(r) \int_s^r \big(f(u) - f(r) \big) |r-u|^{2H-2} \, du \, dr$$
$$- H \int_s^t g(r) f(r) |r-s|^{2H-1} \, dr, \tag{3.13}$$

which is bounded by some multiple of

$$\|f\|_{\mathcal{C}^\beta} \|g\|_\infty |t-s|^{2H+\beta} + \|f\|_\infty \|g\|_\infty |t-s|^{2H}.$$

Combining this with Proposition 3.4, we conclude that

$$\mathbb{Z}_{s,t}^{f,g} = \lim_{\delta \to 0} \mathbb{Z}_{s,t}^{\delta,f,g}, \tag{3.14}$$

exists in probability, is independent of the choice of mollification, and satisfies the bound

$$|\mathbf{E} \mathbb{Z}_{s,t}^{f,g}| \lesssim \|f\|_{\mathcal{C}^\beta} \|g\|_\infty |t-s|^{2H+\beta} + \|f\|_\infty \|g\|_\infty |t-s|^{2H}. \tag{3.15}$$

It now suffices to set $\mathbb{Z}_{s,t}^{f,ij} = \mathbb{Z}_{s,t}^{f_i,f_j}$. Both the fact that Chen's relation holds and the fact that the resulting rough path is geometric follow at once from the fact that these properties hold for the smooth approximations.

Combining (3.15) with the fact that the rough path obtained from [18, Thm 35] satisfies the bound (3.9) with $\alpha = H$ as a consequence of Lemma 3.7, the claim follows.

We now turn to the case $H = \frac{1}{2}$ where it is well known that $\mathbb{Z}_{s,t}^{\delta,f,g}$ defined in (3.11) converges to the Stratonovich integral $\int_s^t \int_s^r f(u) \, dB(u) \, g(r) \circ dB(r)$, so that

$$\mathbb{Z}_{s,t}^{f,g} = \int_s^t \int_s^r f(u) \, dB(u) \, g(r) \, dB(r) + \frac{1}{2} \int_s^t \langle f(u), g(u) \rangle \, du.$$

A simple consequence of Hölder's inequality then leads to the bounds

$$\mathbf{E}|Z_{s,t}^f|^2 \lesssim \|f\|_{L^p} |t-s|^{1-\frac{2}{p}}, \qquad \mathbf{E}|\mathbb{Z}_{s,t}^{f,g}|^2 \lesssim \|f\|_{L^p} \|g\|_{L^p} |t-s|^{2-\frac{4}{p}},$$

where $\|f\|_{L^p} = \|f\|_{L^p([0,1])}$. This shows again that the bound (3.9) holds, this time with $\alpha = \frac{1}{2} - \frac{1}{p}$ (and q arbitrary), and our condition on p guarantees that this is greater than $\frac{1}{3}$.

For $H > \frac{1}{2}$, the first identity in (3.8) above combined with the positivity of the distribution $-\eta''$ and Hölder's inequality yields the bound

$$\mathbf{E}(Z_{s,t}^f)^2 \lesssim |t-s|^{1-\frac{2}{p}} \|f\|_{L^p}^2 \Big| \int_{s-t}^{t-s} \eta''(v) \, dv \Big| \lesssim \|f\|_{L^p}^2 |t-s|^{2H-\frac{2}{p}}.$$

Similarly, again as a consequence of the positivity of $-\eta''$, we have the bound

$$\mathbf{E}(\mathbb{Z}_{s,t}^{f,g})^2 \le \mathbf{E}\big((Z_{s,t}^{|f|})^2 (Z_{s,t}^{|g|})^2 \big) \le 3\mathbf{E}(Z_{s,t}^{|f|})^2 \mathbf{E}(Z_{s,t}^{|g|})^2 \lesssim \|f\|_{L^p}^2 \|g\|_{L^p}^2 |t-s|^{4H-\frac{4}{p}},$$

which shows again that (3.9) holds. $\square$

3.2. Construction and convergence of the rough driver as $\delta \to 0$. The aim of this section is to construct the rough path $\mathbf{Z}^\varepsilon$ (this is the content of Proposition 3.11) and to show that this construction enjoys good stability properties. This is done by stitching together the "canonical" rough path lift for the collection $\{Z^\varepsilon(x)\}_{x \in \mathbf{R}^d}$ obtained in Proposition 3.8. For $H > \frac{1}{2}$, this is just iterated Young integrals. For $H = \frac{1}{2}$ the iterated integrals are considered in the Stratonovich sense and, thanks to the independence of Y and B, the first order process can be interpreted indifferently as either an Itô or a Stratonovich integral.

In order to make use of Proposition 3.8, we use the following lemma, where Y_t^ε denotes the Markov process from Sect. 2.1.

Lemma 3.10. *Let U be as in Proposition 3.8 and let Assumption 2.2 hold for some $p_\star$. When $H < \frac{1}{2}$, we further assume $E \subset L^{p_\star}$ and $\beta < H - \frac{1}{p_\star}$, where $U = \mathcal{C}^\beta$. Then, given $f \in E$ and setting $\hat{f}(t) = f(Y_t^\varepsilon)$, one has for every $p < p_\star$ the bound*

$$\mathbf{E}\|\hat{f}\|_U^p \lesssim \|f\|_E^p,$$

uniformly over E. (Here we use the convention $p_\star = \infty$ when $H > \frac{1}{2}$.)

Proof. For $H \in (\frac{1}{3}, \frac{1}{2})$, the assertion follows immediately from Kolmogorov's continuity test, using Assumption 2.2 and $E \subset L^p$. For $H \geq \frac{1}{2}$, it suffices to note that if $f \in L^p(\mathcal{Y}, \mu)$, then for any fixed $\varepsilon > 0$ the map $\hat{f} \colon t \mapsto f(Y_t^\varepsilon)$ belongs to $L^p([0, 1])$ almost surely and $\mathbf{E}\|\hat{f}\|_{L^p}^p \leq \|f\|_{L^p}^p$. $\qquad\square$

We now show how to collect these objects into one "large" Banach space-valued rough path. The process itself will take values in $\mathcal{B} = \mathcal{C}_b^3(\mathbf{R}^d, \mathbf{R}^d)$, with the second order processes taking value in $\mathcal{B}_2 = \mathcal{C}_b^3(\mathbf{R}^d \times \mathbf{R}^d, (\mathbf{R}^d)^{\otimes 2})$. Our aim is then to define a $\mathcal{B} \oplus \mathcal{B}_2$-valued rough path $(Z^\varepsilon, \mathbb{Z}^\varepsilon)$ which is the canonical lift (in the sense of Proposition 3.8 for any finite collection of x's) of

$$(Z_{s,t}^\varepsilon)(x) = \varepsilon^{\frac{1}{2}-H} \int_s^t F(x, Y_r^\varepsilon)\, dB(r). \tag{3.16}$$

Let $\{f_{i,x}^\varepsilon\}_{i \leq d}$ be the collection of maps from $\mathbf{R}_+$ to $\mathbf{R}^m$ determined by

$$\langle f_{i,x}^\varepsilon(t), e \rangle = (F(x, Y_t^\varepsilon)e)_i, \qquad \forall e \in \mathbf{R}^m.$$

With this notation at hand and recalling the construction of Z^f and $\mathbb{Z}^{f,g}$ as in the proof of Proposition 3.8, it is then natural to look for a $\mathcal{B} \oplus \mathcal{B}_2$-valued rough path $(Z^\varepsilon, \mathbb{Z}^\varepsilon)$ such that, for every $x, \bar{x} \in \mathbf{R}^d$, the identities

$$\left(Z_{s,t}^\varepsilon(x)\right)_i = \varepsilon^{\frac{1}{2}-H} Z_{s,t}^{f_{i,x}^\varepsilon}, \quad \left(\mathbb{Z}_{s,t}^\varepsilon(x, \bar{x})\right)_{ij} \stackrel{\text{def}}{=} \varepsilon^{1-2H} \mathbb{Z}_{s,t}^{f_{i,x}^\varepsilon, f_{j,\bar{x}}^\varepsilon}, \tag{3.17}$$

hold almost surely. (Provided that $F(x, \cdot) \in E$ for every x, the right-hand sides make sense by combining Lemma 3.10 with Proposition 3.8.) We claim that this does indeed define a *bona fide* infinite-dimensional rough path

Recall also that a rough path $\mathbf{Z} = (Z, \mathbb{Z})$ is *weakly geometric* if the identity

$$Z_{s,t} \otimes Z_{s,t} = \mathbb{Z}_{s,t} + \mathbb{Z}_{s,t}^\top, \tag{3.18}$$

holds, where the transposition map $(\cdot)^\top \colon \mathcal{B} \otimes \mathcal{B} \to \mathcal{B} \otimes \mathcal{B}$ swapping the two factors is continuously extended to $\mathcal{B}_2$. With these notations, we have the following.

Proposition 3.11. *Let Assumptions* 2.2, 2.3, *and* 2.7 *hold and let* $\alpha \in (\frac{1}{3}, H - \frac{1}{p_\star})$ *if* $H < \frac{1}{2}$ *and* $\alpha \in (\frac{1}{3}, \frac{1}{2})$ *otherwise. Then, for any* $\varepsilon > 0$, *there exists a random rough path* $\mathbf{Z}^\varepsilon = (Z^\varepsilon, \mathbb{Z}^\varepsilon)$ *in* $\mathscr{C}^\alpha([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ *that is weakly geometric and such that, for every* x, (3.17) *holds almost surely.*

Proof. Since the Chen relations and (3.18) are obviously satisfied for smooth approximations as in (3.11), we only need to show that the analytic constraints hold. In other words, for any fixed $T > 0$ and $\varepsilon > 0$, we look for an almost surely finite random variable C_ε such that

$$\|Z^\varepsilon_{s,t}\|_{\mathcal{B}} \leq C_\varepsilon |t - s|^\alpha, \qquad \|\mathbb{Z}^\varepsilon_{s,t}\|_{\mathcal{B}_2} \leq C_\varepsilon |t - s|^{2\alpha},$$

holds uniformly over all $0 \leq s < t \leq T$. By the Kolmogorov criterion for rough paths [15, Thm 3.1], it suffices to show that, for some $\beta > 0$ and $p \geq 1$ such that $\gamma - \frac{1}{p} > \alpha$, one has the bounds

$$\mathbf{E}\|Z^\varepsilon_{s,t}\|^p_{\mathcal{B}} \leq C_{\varepsilon,p}|t - s|^{p\gamma}, \qquad \mathbf{E}\|\mathbb{Z}^\varepsilon_{s,t}\|^{p/2}_{\mathcal{B}_2} \leq C_{\varepsilon,p}|t - s|^{p\gamma}. \tag{3.19}$$

By Lemma A.1 below, it suffices to show that

$$\sup_{x \in \mathbf{R}^d} (1 + |x|)^{\kappa p} \mathbf{E}|D^\ell Z^\varepsilon_{s,t}(x)|^p \lesssim |t - s|^{p\gamma}, \tag{3.20a}$$

$$\sup_{x, \bar{x} \in \mathbf{R}^d} (1 + |x| + |\bar{x}|)^{\frac{\kappa p}{2}} \mathbf{E}|D^k_x D^\ell_{\bar{x}} \mathbb{Z}^\varepsilon_{s,t}(x, \bar{x})|^{\frac{p}{2}} \lesssim |t - s|^{p\gamma}, \tag{3.20b}$$

for $k + \ell \leq 4$ and some p such that $p > (4d/\kappa) \vee (\gamma - \alpha)^{-1}$.

Since for $\ell \leq 4$ we have $D^\ell_x F^*_i(x, \cdot) \in E$ with $\|D^\ell_x F^*_i(x, \cdot)\|_E \lesssim (1 + |x|)^{-\kappa}$ by Assumption 2.7, it follows immediately from Proposition 3.8 combined with Lemma 3.10 that the bound (3.20a) holds for $\gamma = H$ and $p \leq p_\star$ when $H < \frac{1}{2}$ and for any $\gamma < H$ and $p \geq 1$ when $H \geq \frac{1}{2}$. The bound (3.20b) follows in the same way. (These arguments are somewhat formal, but can readily be justified by taking limits of smooth approximations.) □

In order to prove Proposition 3.1, we make use of the following variant "in probability" of the usual tightness criterion for convergence in law.

Proposition 3.12. *Let* $(\mathcal{Z}, d)$ *be a complete separable metric space and let* $\{L_k : k \in \mathbf{N}\}$ *be a countable collection of continuous maps* $L_k : \mathcal{Z} \to \mathbf{R}$ *that separate elements of* $\mathcal{Z}$ *in the sense that, for every* $x, y \in \mathcal{Z}$ *with* $x \neq y$ *there exist* k *such that* $L_k(x) \neq L_k(y)$.

Let $\{Z_n\}_{n \geq 0}$ *and* Z_∞ *be* $\mathcal{Z}$-*valued random variables such that the collection of their laws is tight and such that* $L_k(Z_n) \to L_k(Z_\infty)$ *in probability for every* k. *Then,* $Z_n \to Z_\infty$ *in probability.*

Proof. Let $\hat{d} : \mathcal{Z}^2 \to \mathbf{R}_+$ be the continuous distance function given by

$$\hat{d}(x, y) = \sum_{k \geq 0} 2^{-k} (1 \wedge |L_k(x) - L_k(y)|),$$

and note first that our assumption implies that $\hat{d}(Z_n, Z_\infty) \to 0$ in probability. Given $\varepsilon > 0$, tightness implies that there exists $K_\varepsilon \subset \mathcal{Z}$ compact such that $\mathbf{P}(Z_n \notin K_\varepsilon) \leq \varepsilon$ for every $n \in \mathbf{N} \cup \{\infty\}$. Furthermore, the set $\{(x, y) \in K_\varepsilon \times K_\varepsilon : d(x, y) \geq \varepsilon\}$ is compact, so that $\hat{d}$ attains its infimum δ on it. Since $\hat{d}$ only vanishes on the diagonal,

one has $\delta > 0$ and, since $\hat{d}(Z_n, Z_\infty) \to 0$ in probability, we can find $N > 0$ such that $\mathbf{P}(\hat{d}(Z_n, Z_\infty) \geq \delta) \leq \varepsilon$ for every $n \geq N$.

It follows that, for every $n \geq N$ one has

$$\mathbf{P}(d(Z_n, Z_\infty) \geq \varepsilon) \leq \mathbf{P}(Z_n \notin K_\varepsilon) + \mathbf{P}(Z_\infty \notin K_\varepsilon) + \mathbf{P}(\hat{d}(Z_n, Z_\infty) \geq \delta) \leq 3\varepsilon,$$

which implies the claim.   □

Regarding tightness itself, the following lemma is a slight variation of well known results.

Lemma 3.13. *Let $\hat{\mathcal{B}} \subset \mathcal{B}$ and $\hat{\mathcal{B}}_2 \subset \mathcal{B}_2$ be compact embeddings of Banach spaces such that $\hat{\mathcal{B}} \otimes \hat{\mathcal{B}} \subset \hat{\mathcal{B}}_2$ with $\|v \otimes w\|_{\hat{\mathcal{B}}_2} \lesssim \|v\|_{\hat{\mathcal{B}}}\|w\|_{\hat{\mathcal{B}}}$. Let $\mathcal{A}$ be a collection of random $\hat{\mathcal{B}} \oplus \hat{\mathcal{B}}_2$-valued rough paths such that for some $\alpha_0 > \frac{1}{3}$ and every $\mathbf{Z} = (Z, \mathbb{Z}) \in \mathcal{A}$,*

$$\mathbf{E}\big(\|Z_{s,t}\|_{\hat{\mathcal{B}}}^p + \|\mathbb{Z}_{s,t}\|_{\hat{\mathcal{B}}_2}^{p/2}\big) \leq |t - s|^{p\alpha_0}, \tag{3.21}$$

for some $p > 3/(3\alpha_0 - 1)$. Then, the laws of the $\mathbf{Z}$'s in $\mathcal{A}$ are tight in $\mathscr{C}^\alpha([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ for any $T > 0$ and $\alpha \in (\frac{1}{3}, \alpha_0 - \frac{1}{p})$.

Proof. Write $\mathcal{G}$ for the metric space given by $\mathcal{B} \oplus \mathcal{B}_2$ endowed with the metric

$$d(a \oplus b, \bar{a} \oplus \bar{b}) = \|\bar{a} - a\|_{\mathcal{B}} \vee \big\|\bar{b} - b - \tfrac{1}{2}(\bar{a} + a) \otimes (\bar{a} - a)\big\|_{\mathcal{B}_2}^{1/2}.$$

Recall then that $\mathscr{C}^\alpha([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ can be identified with the usual space of α-Hölder functions with values in $\mathcal{G}$ by identifying $\mathbf{Z} = (Z, \mathbb{Z})$ with the function $t \mapsto \mathbf{Z}_t \overset{\text{def}}{=} Z_{0,t} \oplus \mathbb{Z}_{0,t}$ and noting that, thanks to Chen's relations,

$$d(\mathbf{Z}_s, \mathbf{Z}_t) = \|Z_{s,t}\|_{\mathcal{B}} \vee \|\mathbb{Z}_{s,t} + \tfrac{1}{2} Z_{s,t} \otimes Z_{s,t}\|_{\mathcal{B}_2}^{1/2}.$$

(See [19, Sec. 7.5] for more details and motivation.) Since d generates the same topology on $\mathcal{G}$ as that given by the Banach space structure of $\mathcal{B} \oplus \mathcal{B}_2$, balls of $\hat{\mathcal{B}} \oplus \hat{\mathcal{B}}_2$ are compact in $\mathcal{G}$. The claim then follows at once from Kolmogorov's continuity test, combined with the fact that, given a compact metric space $(\mathcal{X}, d)$ and a compact subset $\mathcal{K}$ of a Polish space $(\mathcal{Y}, \bar{d})$, the set $\mathcal{C}^\beta(\mathcal{X}, \mathcal{K})$ is compact in $\mathcal{C}^\alpha(\mathcal{X}, \mathcal{Y})$ for any $\beta > \alpha$.   □

Proof of of Proposition 3.1. We apply Proposition 3.12 with the metric space $\mathcal{Z}$ given by $\mathscr{C}^\alpha([0, T], \mathcal{B} \oplus \mathcal{B}_2)$, $Z_n = \mathbf{Z}^{\varepsilon, \delta_n}$ for any given sequence $\delta_n \to 0$, and $Z_\infty = \mathbf{Z}^\varepsilon$ as constructed in Proposition 3.11. The continuous maps L_k appearing in the statement are given by the collection of maps $(Z, \mathbb{Z}) \mapsto Z_t(x)$ and $(Z, \mathbb{Z}) \mapsto \mathbb{Z}_{s,t}(x, \bar{x})$ for a countable dense set of times s and t and elements $x, \bar{x} \in \mathbf{R}^d$.

It follows from (3.20) and Lemma A.1 that the bound (3.21) holds for $\mathbf{Z}^{\varepsilon, \delta}$, uniformly in δ (but with ε fixed), so that the required tightness condition holds by Lemma 3.13. For any fixed $\varepsilon > 0$, the convergences in probability $Z_t^{\varepsilon, \delta}(x)_i \to Z_t^\varepsilon(x)_i$ and $\mathbb{Z}_{s,t}^{\varepsilon, \delta}(x, \bar{x})_{ij} \to \mathbb{Z}_{s,t}^\varepsilon(x, \bar{x})_{ij}$ were shown in Proposition 3.8. (It suffices to apply it with the choices $f = f_{i,x}^\varepsilon$ and $g = f_{j,\bar{x}}^\varepsilon$.)   □

3.3. Formulation of the main technical result. The main technical result of this article can then be formulated in the following way.

Theorem 3.14. *Let $H \in (\frac{1}{3}, 1)$, let Assumptions 2.1–2.3, and 2.7 hold, and let α and $\mathbf{Z}^\varepsilon$ be as in Proposition 3.11. Then, as $\varepsilon \to 0$, $\mathbf{Z}^\varepsilon$ converges weakly in the space of α-Hölder continuous $(\mathcal{B}, \mathcal{B}_2)$-valued rough paths to a limit $\mathbf{Z}$. Furthermore, there is a Gaussian random field W as in (1.8) such that*

$$Z_{s,t}(x) = W(x, t) - W(x, s), \qquad \mathbb{Z}_{s,t}(x, \bar{x}) = \mathbb{W}^{\text{Itô}}_{s,t}(x, \bar{x}) + \Sigma(x, \bar{x})(t - s),$$

where $\mathbb{W}^{\text{Itô}}_{s,t}(x, \bar{x}) = \int_s^t Z_{s,r}(x) \otimes W(\bar{x}, dr)$, interpreted as an Itô integral and Σ is as in (1.6).

The proof of this result will be given in Sects. 4 and 5 below, see Proposition 5.1 which is just a slight reformulation of Theorem 3.14. We first show in Sect. 4 that the family $\{\mathbf{Z}^\varepsilon\}_{\varepsilon \leq 1}$ is tight in a suitable space of rough paths and then identify its limit in Sect. 5.

Corollary 3.15. *Under the assumptions of Theorem 3.14, the solution flow of (1.4) converges weakly to that of the Kunita-type SDE (1.7).*

Proof. Define the $\mathcal{B}$-valued process

$$Z^{0,\varepsilon}_{s,t}(x) = \int_s^t F_0(x, Y^\varepsilon_r) \, dr.$$

It follows from Assumption 2.7 that, for $k \leq 4$ and $p \leq p_\star$,

$$\|D^k Z^{0,\varepsilon}_{s,t}(x)\|_{L^p} \leq \int_s^t \|D^k F_0(x, Y^\varepsilon_r)\|_{L^p} \, dr \lesssim |t - s|(1 + |x|)^{-\kappa},$$

uniformly over ε. This shows that the family $\{Z^{0,\varepsilon}\}_{\varepsilon \leq 1}$ is tight in $\mathcal{C}^\beta([0, 1], \mathcal{B})$ for every $\beta < 1$.

Furthermore, by the ergodic theorem which holds under Assumption 2.1, for every x,

$$\lim_{\varepsilon \to 0} Z^{0,\varepsilon}_{s,t}(x) = (t - s) \int_{\mathcal{Y}} F_0(x, y) \, \mu(dy),$$

almost surely. Since we can choose β and α such that $\alpha + \beta > 1$ and $2\beta > 1$, it follows that there is no need to control any cross terms between $Z^{0,\varepsilon}$ and either Z^ε or $Z^{0,\varepsilon}$ itself in order to be able to solve equations driven by both [27,53]. Furthermore, since the limit of $Z^{0,\varepsilon}$ is deterministic, one deduces joint convergence from Theorem 3.14.

By the continuity theorem for rough differential equations, the solutions of (1.4) written in the form (3.2) converge weakly to those of the rough differential equation

$$dX_t = \delta(X_t) \, d\mathbf{Z}_t + \delta(X_t) \, d\mathbf{Z}^0_t. \tag{3.22}$$

It remains to identify solutions to this equation with those of (1.7).

This is straightforward and follows as in [15, Sec 5.1] for example. Since the Gubinelli derivative x' of the solution $X = (x, x')$ to (3.22) is given by $\delta(x)$, the integral $\int_0^t \delta(X_s) d\mathbf{Z}_s$ is obtained as limit of the compensated Riemann sum

$$\sum_{[u,v] \subset \mathcal{P}} \left(\delta(x_u) Z_{u,v} + (D\delta \cdot \delta)(x_u)(\mathbb{W}_{u,v}^{\text{Itô}} + (u - v)\Sigma) \right)$$

$$= \sum_{[u,v] \subset \mathcal{P}} \left(Z_{u,v}(x_u) + (u - v)G(x_u) \right) + \sum_{[u,v] \subset \mathcal{P}} (D\delta \cdot \delta)(x_u)\mathbb{W}_{u,v}^{\text{Itô}}, \quad (3.23)$$

where $\mathcal{P}$ is a partition on $[0, t]$ and $D\delta \cdot \delta$ is as in (3.3). Since x is continuous and adapted to the filtration generated by W, the first term converges to

$$\int_0^t W(x_u, du) + \int_0^t G(x_u) \, du.$$

The last term on the other hand converges to 0 in probability since it is a discrete martingale and its summands are centred random variables of variance $\mathcal{O}(|v - u|^2)$. $\quad \square$

4. Tightness of the Rough Driver as $\varepsilon \to 0$

The content of this section is the proof of the following tightness result. Let $\{\mathbf{Z}^\varepsilon\}_{\varepsilon \leq 1}$ be given as in Proposition 3.11.

Proposition 4.1. *Let Assumptions 2.1–2.3, and 2.7 hold. For $H \in (\frac{1}{3}, \frac{1}{2}]$, there exists $\alpha \in (\frac{1}{3}, H)$ such that the family $\{\mathbf{Z}^\varepsilon\}_{\varepsilon \leq 1}$ is tight in $\mathscr{C}^\alpha([0, T], \mathcal{B} \oplus \mathcal{B}_2)$.*

For $H \in (\frac{1}{2}, 1)$, if in addition $\int F(x, y) \mu(dy) = 0$ for every x, then the family of rough paths $\mathbf{Z}^\varepsilon$ is tight in $\mathcal{C}^\alpha$ for every $\alpha \in (\frac{1}{3}, \frac{1}{2})$.

It will be convenient to introduce the following notation. Given $f, g \in E$, we use the shorthand

$$J_{s,t}^\varepsilon(f) = \varepsilon^{\frac{1}{2} - H} Z_{s,t}^{f(Y^\varepsilon)}, \qquad \mathbb{J}_{s,t}^\varepsilon(f) = \varepsilon^{1 - 2H} \mathbb{Z}_{s,t}^{f(Y^\varepsilon), g(Y^\varepsilon)}.$$

We then have the following tightness criterion.

Lemma 4.2. *Let $p > d + 1$. Assume that for any $f, g \in E$, $|s - t| \leq 1$, and $\varepsilon \in (0, 1]$,*

$$\|J_{s,t}^\varepsilon(f)\|_{L^p} \leq C\|f\|_E |t - s|^{\alpha_0}, \qquad \left\|\mathbb{J}_{s,t}^\varepsilon(f, g)\right\|_{L^p} \leq C\|f\|_E \|g\|_E |t - s|^{2\alpha_0},$$

where $p > 3/(3\alpha_0 - 1)$. Let furthermore Assumption 2.7 hold with $\kappa > \frac{8d}{p}$. Then the family $\{\mathbf{Z}^\varepsilon\}_{\varepsilon \leq 1}$ is tight in $\mathscr{C}^\alpha([0, T], \mathcal{B} \oplus \mathcal{B}_2)$ for any $\alpha < \alpha_0 - 1/p$.

Proof. Recall that with the above notations, one has from (3.17)

$$\left(Z_{s,t}^\varepsilon(x)\right)_i = J_{s,t}^\varepsilon\left(F_i^*(x, \cdot)\right), \qquad \left(\mathbb{Z}_{s,t}^\varepsilon(x, \bar{x})\right)_{ij} = \mathbb{J}_{s,t}^\varepsilon\left(F_i^*(x, \cdot), F_j^*(\bar{x}, \cdot)\right).$$

By the assumption for $|\ell| \leq 3$ one has for $|s - t| \leq 1$ and $|x - x'| \leq 1$,

$$\|D^\ell Z_{s,t}^\varepsilon(x) - D^\ell Z_{s,t}^\varepsilon(x')\|_{L^p} \lesssim \sum_i \left\| J_{s,t}^\varepsilon\left(D_x^\ell F_i^*(x, \cdot) - D_x^\ell F_i^*(x', \cdot)\right) \right\|_{L^p}$$

$$\lesssim \|D_x^\ell F(x, \cdot) - D_x^\ell F(x', \cdot)\|_E |t - s|^{\alpha_0}$$

$$\lesssim \sup_{y \in [x,x']} \|D_x^\ell F(y, \cdot)\|_E |x - x'||t - s|^{\alpha_0}$$

$$\lesssim (1 + |x|)^{-\kappa} |x - x'||t - s|^{\alpha_0},$$

where we wrote $[x, x']$ for the convex hull of $\{x, x'\}$. Here, the last bound follows from Assumption 2.7. It then follows from Lemma A.1 that, for $\hat{\mathcal{B}}$ as defined in the appendix, $\mathbf{E}\|Z_{s,t}^{\varepsilon}\|_{\hat{\mathcal{B}}}^{p} \lesssim |t-s|^{p\alpha_0}$. We choose ζ to be any number in $(0, 1 - d/p)$.

It similarly follows that for $|k + \ell| \leq 3$ and $|x - x'| \leq 1$

$$
\begin{aligned}
\|D_x^k D_{\bar{x}}^\ell \mathbb{Z}_{s,t}^\varepsilon(x, \bar{x}) &- D_x^k D_{\bar{x}}^\ell \mathbb{Z}_{s,t}^\varepsilon(x', \bar{x})\|_{L^{p/2}} \\
&\lesssim \sum_{i,j} \|\mathbb{J}_{s,t}^\varepsilon \big(D_x^k F_i^*(x, \cdot) - D_x^k F_i^*(x', \cdot), D_x^\ell F_j^*(\bar{x}, \cdot) \big)\|_{L^{p/2}} \\
&\lesssim \sup_{y \in [x, x']} \|D_x^k F(y, \cdot)\|_E \|D_x^\ell F(\bar{x}, \cdot)\|_E |x - x'||t - s|^{2\alpha_0} \\
&\lesssim (1 + |x|)^{-\kappa}(1 + |\bar{x}|)^{-\kappa}|x - x'||t - s|^{2\alpha_0},
\end{aligned}
$$

(and analogously when varying $\bar{x}$). Since $\kappa p / 8 > d$ by assumption, we can again apply Lemma A.1 with $p/2$, thus yielding $\mathbf{E}\|\mathbb{Z}_{s,t}^\varepsilon\|_{\hat{\mathcal{B}}_2}^{p/2} \lesssim |t - s|^{p\alpha_0}$. Since furthermore $p > 3/(3\alpha_0 - 1)$ by assumption, the conditions of Lemma 3.13 are satisfied and the claim follows for any $\alpha < \alpha_0 - 1/p$.   □

Proof of Proposition 4.1. The arguments are quite different for the different ranges of H, but they will always reduce to verifying the assumptions of Lemma 4.2.

First let $H \in (\frac{1}{3}, \frac{1}{2})$. The first assumption of Lemma 4.2 follows from Proposition 4.5 below with $\alpha_0 = H$ and from the trivial bound

$$
\varepsilon^{\frac{1}{2} - \alpha_0} t_0^\alpha \vee \sqrt{t} \lesssim t_0^\alpha, \qquad \forall \varepsilon \leq 1, \quad t \leq T,
$$

while the second assumption follows from Proposition 4.7 below. Both hold for any $p \leq p_\star/4$ where $p_\star > \max\{4d, \frac{6}{(3H-1)}\}$, and the proofs of the propositions are the content of Sect. 4.1.

The ingredients for showing tightness of $\mathbf{Z}^\varepsilon$ where $H \in (\frac{1}{2}, 1)$ are given in Sect. 4.2, starting with a bound on J analogous to that of Proposition 4.5. Unlike in the proof of that statement though, we do *not* show this by bounding the conditional variance $\mathbf{E}(|J_{s,t}^\varepsilon(f)|^2 \,|\, \mathcal{F}^Y)$. This is because, as a consequence of the lack of integrability at infinity of η'' when $H > \frac{1}{2}$, it appears difficult to obtain a sufficiently good bound on it, especially for H close to 1. (In particular, the best bounds one can expect to obtain from a quantitative law of large numbers don't appear to be sufficient when $H > \frac{3}{4}$.) The required bounds are collected in Corollary 4.17 which yields the assumptions of Lemma 4.2 with $\alpha_0 = \frac{1}{2}$ and arbitrary p.

Finally we take $\alpha_0 = \frac{1}{2}$ when $H = \frac{1}{2}$, then $\|J_{s,t}^\varepsilon(f)\|_{L^p} \leq \sqrt{\int_s^t |f(Y_r^\varepsilon)|^2 dr} \lesssim C\|f\|_E \sqrt{|t - s|}$ by Burkholder-Davies-Gundy inequality, and similarly the second order processes satisfies the bound:

$$
\begin{aligned}
\big\|\mathbb{J}_{s,t}^\varepsilon(f, g)\big\|_{L^p} = \Big\| \int_0^t \int_0^u f(Y_v^\varepsilon) g(Y_u^\varepsilon) dB_v dB_u &+ \frac{1}{2} \int_0^t f(Y_r^\varepsilon) g(Y_s^\varepsilon) \Big\|_{L^p} \\
&\lesssim \|f\|_E \|g\|_E |t - s|,
\end{aligned}
$$

allowing us to again apply Lemma 4.2 and concluding the proof.   □

4.1. The low regularity case. This section consists of a number of a priori moment bounds, which we then combine at the end to provide the proof of Proposition 4.7. These uniform in ε moment bounds follow from the Hölder continuity of Y in a subspace of L^{p^*} for a sufficiently large p^*, in particular ergodicity of Y does not play any role.

We will make repeated use of the following simple calculation, where we recall (3.5) for the definition of the distribution η.

Lemma 4.3. *Given $t > 0$ and $H < \frac{1}{2}$, let $\Psi: [0, t]^2 \to \mathbf{R}$ be a continuous function such that for some numbers $\varepsilon > 0$, $\beta > -2H$, $\gamma, \zeta > 1 - 2H$, and $C, \hat{C}, \bar{C} \geq 0$ it holds that*

$$|\Psi(r, r)| \leq C|r|^{\beta}, \quad |\Psi(s, r) - \Psi(r, r)| \leq \hat{C}|r|^{\beta}\left(1 \wedge \tfrac{|s-r|^{\gamma}}{\varepsilon^{\gamma}}\right) + \bar{C}\varepsilon^{\beta-\zeta}|s - r|^{\zeta},$$
(4.1)

for all $s, r \in [0, t]$. Then, one has the bound

$$\left|\int_0^t \int_0^t \Psi(s, r)\eta''(r - s)\, ds\, dr\right| \leq K\left(Ct^{2H+\beta} + \hat{C}t^{\beta+1}\varepsilon^{2H-1} + \bar{C}t^{\zeta+2H}\varepsilon^{\beta-\zeta}\right),$$
(4.2)

with the proportionality constant K depending only on β, γ and ζ. The same bound holds if the upper limit of the inner integral in (4.2) is given by r instead of t.

Proof. Let I be the double integral appearing in (4.2). As a consequence of Lemma 3.6, we can rewrite it as

$$I = -2\alpha_H \int_0^t \int_0^t \left(\Psi(s, r) - \Psi(r, r)\right)|r - s|^{2H-2}\, ds\, dr$$
$$+ 2H \int_0^t \Psi(r, r)\left(|t - r|^{2H-1} + |r|^{2H-1}\right) dr \stackrel{\text{def}}{=} I_1 + I_2.$$

We then have

$$|I_1| \lesssim \hat{C} \int_0^t r^{\beta} \int_0^t \left(1 \wedge (|s - r|/\varepsilon)^{\gamma}\right)|r - s|^{2H-2}\, ds\, dr$$
$$+ \bar{C}\varepsilon^{\beta-\zeta} \int_0^t \int_0^t |r - s|^{2H+\zeta-2}\, ds\, dr$$
$$\lesssim \hat{C}\varepsilon^{2H-1} \int_0^t r^{\beta} \int_{\mathbf{R}} (1 \wedge |u|^{\gamma})|u|^{2H-2}\, du\, dr + \bar{C}\varepsilon^{\beta-\zeta} t^{2H+\zeta}$$
$$\lesssim \hat{C}\varepsilon^{2H-1}t^{\beta+1} + \bar{C}\varepsilon^{\beta-\zeta}t^{2H+\zeta}.$$

We used the conditioned imposed on β, γ and ζ. Regarding I_2, we have the bound

$$|I_2| \lesssim C \int_0^t |r|^{\beta}\left(|t - r|^{2H-1} + |r|^{2H-1}\right) dr$$
$$= Ct^{2H+\beta} \int_0^1 |r|^{\beta}\left(|1 - r|^{2H-1} + |r|^{2H-1}\right) dr,$$

and the claim follows. □

Note that replacing $\Psi(r, s)$ by $\Psi_\tau(r, s) = \tau^{2H}\Psi(\tau r, \tau s)$ and t by t/τ, the left-hand side of (4.2) is left unchanged. Regarding the bounds (4.1), such a change leads to the substitutions $C \mapsto C\tau^{2H+\beta}$, $\varepsilon \mapsto \varepsilon/\tau$, $\hat{C} \mapsto \hat{C}\tau^{2H+\beta}$, and $\bar{C} \mapsto \bar{C}\tau^{2H+\beta}$. All three terms appearing in the right-hand side of (4.2) are invariant under these substitutions.

Remark 4.4. The proof of Lemma 4.3 works *mutatis mutandis* for Ψ taking values in a Banach space, for example L^p. We also see that if Ψ is upper bounded by a finite sum of terms of the type (4.1) with different exponents β and γ, then the bound (4.2) still holds with the corresponding sum in the right-hand side.

We perform a number of preliminary calculations. For this, it will be notationally convenient to introduce the shortcuts

$$I^\varepsilon_{s,t}(f) = \int_s^t f(Y_r^\varepsilon)\, dB_r, \qquad J^\varepsilon_{s,t}(f) = Z^{f(Y_\cdot^\varepsilon)}_{s,t} = \varepsilon^{\frac{1}{2}-H} I^\varepsilon_{s,t}(f),$$

for $f \in E$ (with values in $\mathbf{R}^m$).

Proposition 4.5. *Let* $H \in (\frac{1}{3}, \frac{1}{2})$ *and let Assumptions* 2.2 *and* 2.3 *hold for some* $p_\star \geq 2$. *Then there exists a constant* C *such that, uniformly over* $s \geq 0$, $t \geq 0$, *and* $f \in E$,

$$\|J^\varepsilon_{s,t}(f)\|_{L^{p_\star}} \leq C\|f\|_E\big(\varepsilon^{\frac{1}{2}-H}|t-s|^H \vee \sqrt{|t-s|}\big). \tag{4.3}$$

Proof. Let $\mathcal{F}^Y$ denote the σ-algebra generated by all point evaluations of the process Y, and $\mathcal{F}^Y_t$ the corresponding filtration. Write p for $p_\star$ for brevity. Since B is independent of $\mathcal{F}^Y$ and the L^p norm of an element of a Wiener chaos of fixed degree is controlled by its L^2 norm, we have

$$\|I^\varepsilon_{s,t}(f)\|^2_{L^p} = \big|\mathbf{E}\big(\mathbf{E}\big(I^\varepsilon_{s,t}(f)^p \mid \mathcal{F}^Y\big)\big)\big|^{2/p} \leq c\|\mathbf{E}(I^\varepsilon_{s,t}(f)^2 \mid \mathcal{F}^Y)\|_{L^{p/2}}, \tag{4.4}$$

for some universal constant c depending only on p and on the degree of the Wiener chaos, so that

$$\|I^\varepsilon_{s,t}(f)\|^2_{L^p} \lesssim \Big\|\int_s^t \int_s^t f(Y_r^\varepsilon) f(Y_{r'}^\varepsilon) \eta''(r-r')\, dr\, dr'\Big\|_{L^{p/2}}. \tag{4.5}$$

Since $f \in E$ is in L^p by Assumption 2.3, it follows from Assumption 2.2 that

$$\big\|f(Y_u^\varepsilon) f(Y_{u+v}^\varepsilon) - f(Y_u^\varepsilon)^2\big\|_{L^{p/2}} \lesssim \|f\|^2_E\big(|v/\varepsilon|^H \wedge 1\big).$$

We can therefore apply Lemma 4.3 with $\gamma = H$ and $\beta = 0$ so that, for $\|f\|_E \leq 1$, one has

$$\|I^\varepsilon_{s,t}(f)\|^2_{L^p} \lesssim \varepsilon^{2H-1}|t-s| + |t-s|^{2H}, \tag{4.6}$$

whence the desired bound follows. (The condition $\gamma > 1 - 2H$ is satisfied since $H > \frac{1}{3}$ by assumption.) $\square$

We now consider the second-order process $\mathbb{J}$ given by

$$\mathbb{J}^\varepsilon_{s,t}(f, g) = \varepsilon^{1-2H} \mathbb{Z}^{f(Y_\cdot^\varepsilon), g(Y_\cdot^\varepsilon)}_{s,t} = \varepsilon^{1-2H} \int_s^t \int_s^v f(Y_u^\varepsilon)\, dB(u)\, g(Y_v^\varepsilon)\, dB(v), \tag{4.7}$$

and bound it in a similar way. Recalling that $\langle f, g\rangle_\mu = \int_\mathcal{Y}\langle f, g\rangle d\mu$, we first obtain a bound on its expectation.

Proposition 4.6. *Let* $H \in (\frac{1}{3}, \frac{1}{2})$, *let Assumptions* 2.2 *and* 2.3 *hold for some* $p_\star \geq 2$, *and let* $f, g \in E$. *One has*

$$\left\|\mathbf{E}\big(\mathbb{J}_{s,t}^\varepsilon(f, g) \,|\, \mathcal{F}^Y\big)\right\|_{L^p} \lesssim \|f\|_E \|g\|_E \big(\varepsilon^{1-2H} |t - s|^{2H} \vee |t - s|\big),$$

provided that $2p \leq p_\star$.

Proof. It follows from (4.7) that we have the identity

$$\mathbf{E}\big(\mathbb{J}_{s,t}^\varepsilon(f, g) \,|\, \mathcal{F}^Y\big) = \frac{\varepsilon^{1-2H}}{2} \int_s^t \int_s^v f(Y_u^\varepsilon) g(Y_v^\varepsilon) \, \eta''(u - v) \, du \, dv,$$

and we conclude from Lemma 4.3 and the bound $\left\|g(Y_u^\varepsilon)(f(Y_{u+w}^\varepsilon) - f(Y_u^\varepsilon))\right\|_{L^p} \lesssim \|f\|_E \|g\|_E \big(|w/\varepsilon|^H \wedge 1\big)$ exactly as above. $\square$

Proposition 4.7. *Let* $H \in (\frac{1}{3}, \frac{1}{2})$, *let Assumptions* 2.2 *and* 2.3 *hold for some* $p_\star \geq 2$, *let* $f, g \in E$, *and let* $2p \leq p_\star$. *Then there exists a constant* C *such that*

$$\left\|\mathbb{J}_{s,t}^\varepsilon(f, g)\right\|_{L^p} \leq C \|f\|_E \|g\|_E \big(\varepsilon^{1-2H} |t - s|^{2H} \vee |t - s|\big).$$

If g *is a constant, one obtains a stronger upper bound of the form*

$$\|f\|_E |g| \big(\varepsilon^{\frac{1}{2}-H} |t - s|^{\frac{1}{2}+H} \vee \varepsilon^{1-2H} |t - s|^{2H}\big).$$

Proof. By Proposition 4.6, it suffices to obtain a bound on

$$\left\|\mathbb{J}_{s,t}^\varepsilon(f, g) - \mathbf{E}\big(\mathbb{J}_{s,t}^\varepsilon(f, g) \,|\, \mathcal{F}^Y\big)\right\|_{L^p}.$$

As a consequence of Proposition 3.4 and (4.4), we have the bound

$$\left\|\mathbb{J}_{s,t}^\varepsilon(f, g) - \mathbf{E}\big(\mathbb{J}_{s,t}^\varepsilon(f, g) \,|\, \mathcal{F}^Y\big)\right\|_{L^p} \leq \sqrt{2} \left\|\tilde{\mathbb{J}}_{s,t}^\varepsilon(f, g)\right\|_{L^p},$$

where we set

$$\tilde{\mathbb{J}}_{s,t}^\varepsilon(f, g) = \varepsilon^{1-2H} \int_s^t \int_s^v f(Y_u^\varepsilon) \, dB(u) \, g(Y_v^\varepsilon) \, d\tilde{B}(v),$$

for a fractional Brownian motion $\tilde{B}$ independent of B (and Y). We furthermore restrict ourselves to the case $s = 0$ and $m = 1$ without loss of generality.

At this point we note that for every $H > \frac{1}{3}$, one has the identity

$$\mathbf{E}\left(\big(\tilde{\mathbb{J}}_{0,t}^\varepsilon(f, g)\big)^2 \,\Big|\, \mathcal{F}^Y\right) = \frac{1}{2} \int_0^t \int_0^t \phi_\varepsilon(s, s') \eta''(s - s') \, ds \, ds', \tag{4.8}$$

where we have set

$$\phi_\varepsilon(s, s') = \frac{\varepsilon^{2-4H}}{2} g(Y_s^\varepsilon) g(Y_{s'}^\varepsilon) \int_0^s \int_0^{s'} f(Y_r^\varepsilon) f(Y_{r'}^\varepsilon) \, \eta''(r - r') \, dr \, dr'.$$

As a consequence of (4.4), we deduce from (4.8) the bound

$$\|\tilde{\mathbb{J}}_{0,t}^\varepsilon(f, g)\|_{L^p}^2 \lesssim \left\| \int_0^t \int_0^t \phi_\varepsilon(s, s') \eta''(s - s') \, ds \, ds' \right\|_{L^{p/2}}. \tag{4.9}$$

We now bound ϕ_ε in such a way that Lemma 4.3 (combined with Remark 4.4) applies with $\Psi = \phi_\varepsilon$. In order to apply this result, we first verify that the first bound in (4.1) holds. Applying Hölder's inequality we obtain for $2p \le p_\star$,

$$\left\| \phi_\varepsilon(s,s) \right\|_{L^{p/2}} \le C\varepsilon^{2-4H} \| g(Y_s^\varepsilon) \|_{L^{2p}}^2 \left\| \int_0^s \int_0^s f(Y_r^\varepsilon) f(Y_{r'}^\varepsilon) \eta''(r-r') \, dr \, dr' \right\|_{L^p}.$$

(4.10)

Since the last factor is the same expression as the right-hand side of (4.5), it is bounded as in (4.6), thus yielding

$$\left\| \phi_\varepsilon(s,s) \right\|_{L^{p/2}} \lesssim \varepsilon^{2-4H} \| f \|_E^2 \| g \|_E^2 \left(s^{2H} \vee \varepsilon^{2H-1} s \right).$$

Regarding the second bound in (4.1), we note that, for $s' \ge s$ and $\alpha > \frac{1}{3}$, one has

$$\phi_\varepsilon(s,s') - \phi_\varepsilon(s,s) = \varepsilon^{2-4H} g(Y_s^\varepsilon) g(Y_{s'}^\varepsilon) \int_0^s \int_s^{s'} f(Y_r^\varepsilon) f(Y_{r'}^\varepsilon) |r-r'|^{2H-2} \, dr \, dr'$$
$$+ \varepsilon^{2-4H} g(Y_s^\varepsilon) \big(g(Y_{s'}^\varepsilon) - g(Y_s^\varepsilon) \big) \int_0^s \int_0^s f(Y_r^\varepsilon) f(Y_{r'}^\varepsilon) \eta''(r-r') \, dr \, dr'.$$

Since $2p \le p_\star$ and $\int_0^s \int_s^{s'} |r-r'|^{2H-2} \, dr \, dr' \lesssim |s'-s|^{2H}$, the $L^{p/2}$ norm of the first term is of order $\varepsilon^{2-4H} \| f \|_E^2 \| g \|_E^2 |s'-s|^{2H}$. By Hölder's inequality, the second term is bounded similarly to before by

$$\varepsilon^{2-4H} \| g(Y_s^\varepsilon) \|_{L^{2p}} \| g(Y_s^\varepsilon) - g(Y_{s'}^\varepsilon) \|_{L^{2p}} \left\| \int_0^s \int_0^s f(Y_r^\varepsilon) f(Y_{r'}^\varepsilon) \eta''(r-r') \, dr \, dr' \right\|_{L^p}.$$

By Assumptions 2.2 and 2.3, the factors involving g are bounded by

$$\| g \|_E^2 \left(|s'-s|^H \varepsilon^{-H} \wedge 1 \right),$$

while the remaining factor is the same is in (4.10), thus yielding a bound of the order

$$\| \phi_\varepsilon(s,s') - \phi_\varepsilon(s,s) \|_{L^{p/2}} \lesssim \varepsilon^{2-4H} \| f \|_E^2 \| g \|_E^2 \left(s^{2H} \vee \varepsilon^{2H-1} s \right) \left(|s'-s|^H \varepsilon^{-H} \wedge 1 \right).$$

Applying Lemma 4.3 (and Remark 4.4) and inserting the resulting bound into (4.9), eventually yields the bound

$$\| \tilde{\mathbb{J}}_{0,t}^\varepsilon(f,g) \|_{L^p}^2 \lesssim \| f \|_E^2 \| g \|_E^2 \left(\varepsilon^{2-4H} |t|^{4H} + \varepsilon^{1-2H} |t|^{1+2H} + |t|^2 \right), \qquad (4.11)$$

as desired. (Note that the second term is bounded by the first and the last one which is why it was omitted in the statement.)

In case g is a constant, the second term in the expression for $\phi_\varepsilon(s,s') - \phi_\varepsilon(s,s)$ vanishes identically. Since this is the term responsible for the summand proportional to $|t|^2$ in (4.11), the claim follows. $\square$

4.2. The regular case $H \in (\frac{1}{2}, 1)$. For the case where the slow variables are driven by a fractional Brownian motion of higher regularity, $H > \frac{1}{2}$, we exploit the ergodicity of the fast motion even for proving tightness for the first order processes.

To prove the tightness of the processes $\mathbf{Z}_t^\varepsilon$, we take a different strategy and estimate higher order moments of the $Z_{s,t}^\varepsilon$ and $\mathbb{Z}_{s,t}^\varepsilon$. This requires us to estimate the expectation of multiple integrals of the form

$$\int_0^t \cdots \int_0^t \prod_{i=1}^{2p} f_i(Y_{t_i}^\varepsilon) \, dB_{t_1} \cdots dB_{t_{2p}}. \tag{4.12}$$

For the second order processes, half of the upper limits of the integrals are given by one of the t_i's, but since we will not need to exploit any cancellations these integrals are controlled by the bounds on the hypercube. For $p = 1$, it is easy to see that this integral is of order $\varepsilon^{2H-1}t$, but the case $p = 2$ is already more complicated:

$$\mathbf{E} \int_0^t \cdots \int_0^t \prod_{i=1}^4 f_i(Y_{t_i}^\varepsilon) \, dB_{t_1} \cdots dB_{t_4}$$

$$= C \int_0^t \cdots \int_0^t \mathbf{E}\Big(\prod_{i=1}^4 f_i(Y_{t_i}^\varepsilon) \Big) |t_2 - t_1|^{2H-2} |t_4 - t_3|^{2H-2} \, dt_1 \cdots dt_4. \tag{4.13}$$

If we look at the regime $t_1 < t_2 < t_3 < t_4$ say and write $P_t^\varepsilon = P_{t/\varepsilon}$, the first factor is given by

$$\mathbf{E}\Big(\prod_{i=1}^4 f_i(Y_{t_i}^\varepsilon) \Big) = \int f_1 P_{t_2-t_1}^\varepsilon \big(f_2 P_{t_3-t_2}^\varepsilon (f_3 P_{t_4-t_3}^\varepsilon f_4) \big) \, d\mu.$$

Since $f_{3,4} = f_3 P_{t_4-t_3}^\varepsilon f_4$ is no longer centred, we unfortunately do not have very good bounds on this expression. One can however do better than $\exp(-c|t_4-t_3|/\varepsilon)$: subtracting and adding the mean of $f_{3,4}$, we can write the expression as

$$\int f_1 P_{t_2-t_1}^\varepsilon \big(f_2 P_{t_3-t_2}^\varepsilon (f_{3,4} - \bar{f}_{3,4}) \big) \, d\mu + \bar{f}_{3,4} \, \bar{f}_{1,2}, \tag{4.14}$$

where now the first term is bounded by $\exp(-c|t_4-t_2|/\varepsilon)$ and the second term is bounded by $\exp(-c|t_4 - t_3|/\varepsilon - c|t_2 - t_1|/\varepsilon)$. This is still not optimal: we note this time that we can recenter $f_{2,3,4} = f_2 P_{t_3-t_2}^\varepsilon (f_{3,4} - \bar{f}_{3,4})$ "for free" since f_1 has mean zero, so the first term is actually of order $\exp(-c|t_4 - t_1|/\varepsilon)$. It is then not too difficult to see that, the contribution of the second term of (4.14) to the integral (4.13) is of order $\varepsilon^{4H-2}t^2$, while the contribution of the first term is $\varepsilon^{4H-1}t$, which is of lower order for $t \geq \varepsilon$. Our aim is to generalise such considerations to arbitrarily high moments.

In particular, the "correct" way of rewriting the factor $\mathbf{E}\big(\prod_{i=1}^{2p} f_i(Y_{t_i}^\varepsilon) \big)$ so that it yields usable bounds is in terms of its cumulants. Given a collection $\{X_i\}_{i \in I}$ of random variables and a subset $A \subset I$, we write X_A as a shorthand for the collection $\{X_i\}_{i \in A}$ and X^A as a shorthand for $\prod_{i \in A} X_i$. Given a finite set A, we write $\mathcal{P}(A)$ for the set of partitions of A. We also write $\mathbf{E}_c X_A$ for the joint cumulant, so that one has the identities

$$\mathbf{E} X^I = \sum_{\Delta \in \mathcal{P}(I)} \prod_{A \in \Delta} \mathbf{E}_c X_A, \qquad \mathbf{E}_c X_I = \sum_{\Delta \in \mathcal{P}(I)} C_\Delta \prod_{A \in \Delta} \mathbf{E} X^A, \tag{4.15}$$

where $C_\Delta = (|\Delta| - 1)!(-1)^{|\Delta|-1}$.

Proposition 4.8. *Let Assumptions* 2.1 *and* 2.3 *hold for $H > \frac{1}{2}$. For any $k \geq 2$ and any $s_1, \ldots, s_k \in [0, t]$, there exist constants $c, C > 0$ such that the following holds. Let $f_1, \ldots, f_k \in E$ with $\int_{\mathcal{Y}} f_i \, d\mu = 0$, let $s_1 < \ldots < s_k$, and set $X_i = f_i(Y_{s_i})$. Then, one has the bound*

$$\mathbf{E}_c X_{[k]} \leq C \exp\left(-c \sum_{i,j \leq k} |s_i - s_j|\right) \prod_i \|f_i\|_E$$

where $[k]$ denotes the set $\{1, \ldots, k\}$.

Proof. Note first that since c is allowed to depend on k, it actually suffices to show that $\mathbf{E}_c X_{[k]} \leq C \exp\left(-c \sup_{i < k} |s_{i+1} - s_i|\right)$. From now on we fix $i_\star \in \{1, \ldots, k\}$ to be the index which realises that supremum. Let $\tilde{Y}$ be an independent copy of Y and set

$$\tilde{X}_j = \begin{cases} f_j(Y_{s_j}) & \text{if } j \leq i_\star, \\ f_j(\tilde{Y}_{s_j}) & \text{otherwise.} \end{cases}$$

The most important property of the joint cumulant of a collection of random variables is that if it can be broken into two independent sub-collections, then the joint cumulant vanishes. As a consequence, we have

$$\mathbf{E}_c X_{[k]} = \mathbf{E}_c X_{[k]} - \mathbf{E}_c \tilde{X}_{[k]} = \sum_{\Delta \in \mathcal{P}(I)} C_\Delta \left(\prod_{A \in \Delta} \mathbf{E} X^A - \prod_{A \in \Delta} \mathbf{E} \tilde{X}^A\right).$$

We now put a total order on the elements of a partition Δ by postulating that $A_1 \leq A_2$ whenever $\inf\{a \in A_1\} \leq \inf\{a \in A_2\}$ (this is just for definiteness, the actual choice of order is unimportant). We can then write the above as a telescoping sum, yielding

$$\mathbf{E}_c X_{[k]} = \sum_{\Delta \in \mathcal{P}(I)} C_\Delta \sum_{A \in \Delta} \left(\mathbf{E} X^A - \mathbf{E} \tilde{X}^A\right) \left(\prod_{B < A, \, B \in \Delta} \mathbf{E} X^B\right) \left(\prod_{B > A, \, B \in \Delta} \mathbf{E} \tilde{X}^B\right).$$

$$(4.16)$$

We fix $A \subset [k]$ such that $\mathbf{E} X^A \neq \mathbf{E} \tilde{X}^A$ and write $A = \{a_1, \ldots, a_\ell\}$ with $\ell = |A|$ and $i \mapsto a_i$ increasing. We also write $j_\star < \ell$ for the index such that $a_{j_\star} \leq i_\star$ and $a_{j_\star+1} > i_\star$. (This necessarily exists since otherwise $\mathbf{E} X^A = \mathbf{E} \tilde{X}^A$.) For $i < \ell$ and $n \geq 1$, we also write $T_i \colon E_n \to E_{n+1}$ for the operator given by

$$T_i g = f_{a_i} P_{t_i} g, \qquad t_i \overset{\text{def}}{=} s_{a_{i+1}} - s_{a_i},$$

whose norm, as an operator from E_n to E_{n+1}, is bounded by a (possibly n-dependent) multiple of $\|f_{a_i}\|_E$, since it is of order $\|f_{a_i}\|_E \, e^{-ct_i}$ from $E \times E_n$ to E_{n+1}, when restricted to functions of vanishing mean, by Assumption 2.1. We used the continuity of multiplication of functions. It then follows from the Markov property that

$$\mathbf{E} X^A = \int_{\mathcal{Y}} T_1 \ldots T_{\ell-1} f_{a_\ell} \, d\mu, \qquad (4.17)$$

(this is easily shown by induction over ℓ) while we similarly have by the definition of $\tilde{X}$

$$\mathbf{E} \tilde{X}^A = \int_{\mathcal{Y}} T_1 \ldots T_{i_\star-1} f_{a_{i_\star}} \, d\mu \int_{\mathcal{Y}} T_{i_\star+1} \ldots T_{\ell-1} f_{a_\ell} \, d\mu. \qquad (4.18)$$
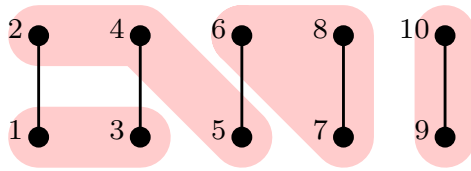
Fig. 1. Example of pairing and partition of 10 elements

This is because, setting $A_1 = \{a_1, \ldots, a_{i_\star}\}$ and $A_2 = A \backslash A_1$, one has $\mathbf{E}\tilde{X}^A = \mathbf{E}X^{A_1}\mathbf{E}X^{A_2}$ by the definition of $\tilde{X}$, so that (4.18) follows from (4.17). Writing $g = T_{i_\star+1} \ldots T_{\ell-1} f_{a_\ell} \in E_{\ell-i_\star}$, it follows that

$$\mathbf{E}X^A - \mathbf{E}\tilde{X}^A = \int T_1 \ldots T_{i_\star}\left(g - \int g\, d\mu\right) d\mu.$$

The spectral gap assumption (2.1) and the definition of $i_\star$ then imply that

$$\left|\mathbf{E}X^A - \mathbf{E}\tilde{X}^A\right| \le C\exp(-c|s_{i_\star+1} - s_{i_\star}|)\prod_{a \in A}\|f_a\|_E.$$

Combining this with (4.16) immediately leads to the claimed bound on the corresponding cumulant. □

The first identity of (4.15) combined with Wick's formula for the moments of Gaussians now suggest that we should rewrite the expectation of (4.12) as a sum over terms indexed by pairs (Δ, π) where Δ is a partition of $[2p]$ arising from (4.15) and representing a product of cumulants of the $f(Y_{t_i})$ and π is a pairing of $[2p]$ arising from Wick's formula.

Figure 1 for example represents the pairing π and partition Δ given by

$$\begin{aligned}\pi &= \{(1,2), (3,4), (5,6), (7,8), (9,10)\},\\ \Delta &= \{\{1,3\}, \{2,4,5\}, \{6,7,8\}, \{9,10\}\}\end{aligned} \tag{4.19}$$

Each pairing (i, j) yields a factor $|s_i - s_j|^{2H-2}$ while each element B of the partition yields an exponential factor of the form $\prod_{i,j \in B}\exp\left(-c|s_i - s_j|\right)$ thanks to Proposition 4.8. Since we consider the case $H > \frac{1}{2}$, this yields a locally integrable function in the expression for the expectation of (4.12), so our analysis mainly focuses on the large-scale behaviour. We will show then that the terms with $\Delta = \pi$, yield a contribution of order $\varepsilon^{(2H-1)p}t^p$ which dominates our bound, while all other terms are of higher order in ε/t. We now proceed to formalising this.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with vertex set $\mathcal{V}$ and edge multiset $\mathcal{E}$ (multiple edges are allowed). Edges $e \in \mathcal{E}$ are oriented from e_- to e_+ and we only consider graphs with $e_+ \ne e_-$. We also label the vertices by two exponents $\alpha_\pm\colon \mathcal{E} \to \mathbf{R}_-$. Finally, we assume that we have a "kernel assignment", i.e. a collection of functions $K_e\colon \mathbf{R} \to \mathbf{R}$ (with $e \in \mathcal{E}$) such that

$$|K_e(t)| \le C\left(|t|^{\alpha_-(e)}\mathbf{1}_{|t|\le 1} + |t|^{\alpha_+(e)}\mathbf{1}_{|t|\ge 1}\right). \tag{4.20}$$

We denote by $\|K_e\|_e$ the smallest possible constant C appearing in the above expression. For those who are not familiar with such graph representations, such a graph will be used to encode expressions of the type

$$I_{\mathcal{G}} := \int_{\mathbf{R}^{nd}} \prod_{e \in \mathcal{E}} K_e(s_{e_+} - s_{e_-})\, \varphi(x_1, \ldots, x_k) dx. \tag{4.21}$$

Each of its nodes u represents an integration variable s_u, each edge e represents a factor K_e, and the resulting expression is integrated against some bounded function φ. The exponents $\alpha_-(e)$ and $\alpha_+(e)$ indicate the singularity of K_e at 0 and at infinity respectively.

Definition 4.9. We say that such a labelled graph is "regular" if, for every subset $\mathcal{V}_0 \subset \mathcal{V}$, we have

$$\sum_{e \in \mathcal{E}\,:\,e_\pm \in \mathcal{V}_0} \alpha_-(e) + |\mathcal{V}_0| > 1. \tag{4.22}$$

The significance of this condition (also called Weinberg's condition) is that it guarantees that the function $K^{\mathcal{G}}$ on $\mathbf{R}^{\mathcal{V}}$ given by

$$K^{\mathcal{G}}(s) = \prod_{e \in \mathcal{E}} K_e(s_{e_+} - s_{e_-}) \tag{4.23}$$

is locally integrable [71] where $s_{e_\pm}$ is the $e_\pm$ component of $s \in \mathbf{R}^{\mathcal{V}}$. (See also [28, Prop. 2.3] for a formulation closer to the one given here.)

   We will be mainly interested in the large-scale behaviour here. To describe this, consider a partition $\mathcal{P}$ of $\mathcal{V}$. We say that such a partition is *tight* if there exists $A \in \mathcal{P}$ such that $A \cap \mathcal{V}_i \neq \emptyset$ for every connected component $\mathcal{V}_i$ of $\mathcal{G}$. Given $\mathcal{P}$, we then also write $u \sim v$ if there exists $A \in \mathcal{P}$ with $\{u, v\} \subset A$.

Definition 4.10. We then say that a labelled graph as above is "integrable" if

$$\sum_{e \in \mathcal{E}\,:\,e_- \nsim e_+} \alpha_+(e) + |\mathcal{P}| < 1, \tag{4.24}$$

for every tight partition $\mathcal{P}$. (Note the similarity with Weinberg's condition.)

The following is then an immediate consequence of [28, Thm 4.3].

Proposition 4.11. *Let $\mathcal{G}$ be a regular and integrable graph with m connected components. Then, there exists C depending on $\mathcal{G}$ such that*

$$\int_{[-L,L]^{\mathcal{V}}} |K^{\mathcal{G}}(s)|\, ds \lesssim L^m \prod_{e \in \mathcal{E}} \|K_e\|_e,$$

uniformly over $L \geq 1$, with proportionality constant depending only on the labelled graph $\mathcal{G}$, where $\|K_e\|_e$ is as defined by (4.20).

Remark 4.12. Our bound on the large-scale behaviour of the kernels K_t is weaker than the bound [28, Eq. 4.1] since we assume no bounds on the derivatives. The reason why the result still holds is that we assume local integrability, which avoids all renormalisation issues and therefore gets rid of regularity requirements.

An immediate, but very useful, corollary is the following.

Corollary 4.13. *Let $\mathcal{G}$ be a regular graph with m connected components and let $L \geq 1$. Let $\beta \colon \mathcal{E} \to \mathbf{R}_+$ be such that*

$$\sum_{e \in \mathcal{E} \,:\, e_- \not\sim e_+} \big(\alpha_+(e) - \beta(e)\big) + |\mathcal{P}| < 1, \tag{4.25}$$

for every tight partition $\mathcal{P}$. Then, there exists C depending on $\mathcal{G}$ and β such that

$$\int_{[-L,L]^{\mathcal{V}}} |K^{\mathcal{G}}(t)|\, dt \leq C L^m \prod_{e \in \mathcal{E}} L^{\beta(e)} \|K_e\|_e.$$

Proof. It suffices to note that we can assume that the kernels K_e vanish outside of $[-2L, 2L]$ since this does not affect the value of the integral. If we then consider the graph identical to $\mathcal{G}$, but with its labels replaced by $(\alpha_-, \alpha_+ - \beta)$, then (4.25) implies integrability for the new graph by (4.24). The local integrability condition (regularity) still holds, so Proposition 4.11 applies. It remains to note that since $1 \leq (L/|t|)^{\beta} \mathbf{1}_{1 \leq |t| \leq L}$, decreasing $\alpha_+(e)$ by $\beta(e)$ in (4.20) has the effect of increasing the norm $\|\cdot\|_e$ by (at most) a factor $(2L)^{\beta(e)}$, provided that we do consider functions supported in $[-2L, 2L]$. □

We will make use of the following property.

Lemma 4.14. *Let $\tilde{\mathcal{G}}$ be a graph obtained by deleting some of the edges of $\mathcal{G}$ but without changing its connected components. If $\tilde{\mathcal{G}}$ is integrable, then so is $\mathcal{G}$ itself.*

Proof. This is immediate from Definition 4.10, combined with the fact that the $\alpha_+(e)$ are negative by assumption. □

The following simple result will also be useful.

Lemma 4.15. *If $\alpha_+(e) < -1$ for every edge e of $\mathcal{G}$, then it is integrable.*

Proof. Let $\mathcal{P}$ be a tight partition of the vertex set $\mathcal{V}$ of $\mathcal{G}$ and let $\mathcal{G}_{\mathcal{P}} = (\mathcal{V}_{\mathcal{P}}, \mathcal{E}_{\mathcal{P}})$ denote the graph obtained by removing self-loops from $\mathcal{G}/\sim$, with $\sim$ obtained from $\mathcal{P}$ as in (4.24). Then $\mathcal{G}_{\mathcal{P}}$ is connected by the definition of tightness so that $|\mathcal{E}_{\mathcal{P}}| \geq |\mathcal{V}_{\mathcal{P}}| - 1$, which translates into $|\{e \in \mathcal{E} \,:\, e_- \not\sim e_+\}| \geq |\mathcal{P}| - 1$. Since $\alpha_+(e) < -1$ for every edge, the bound (4.24), and therefore the desired claim, then follow at once. □

We now use these preliminary results both to bound J and $\mathbb{J}$ and to determine their limits in the case $H > \frac{1}{2}$.

Our main technical result is the following bound.

Proposition 4.16. *Let Assumptions 2.1 and 2.3 hold for $H > \frac{1}{2}$ and let $\kappa \in (0, 2-2H)$. For $f, g \in E$ with $\int f\, d\mu = \int g\, d\mu = 0$, set*

$$C(f, g) = \Gamma(1 - 2H)\big(\langle f, \mathcal{L}^{1-2H} g \rangle_\mu + \langle g, \mathcal{L}^{1-2H} f \rangle_\mu\big).$$

Then, for every $p \geq 1$ and $f \in E^{2p}$ with $\int f_i\, d\mu = 0$ for every i, setting

$$I_{2p}(f) = \int_{L_1}^{M_1} \cdots \int_{L_{2p}}^{M_{2p}} \Big(\prod_{j=1}^{2p} f_j(Y_{t_j}) \Big) \Big(\prod_{k=1}^{p} |t_{2k} - t_{2k-1}|^{2H-2} \Big) dt_1 \cdots dt_{2p},$$

there exists a constant $K > 0$ such that

$$\left| I_{2p}(f) - \prod_{k=1}^{p} |[L_{2k-1}, M_{2k-1}] \cap [L_{2k}, M_{2k}]| C(f_{2k}, f_{2k-1}) \right| \leq K L^{p-\kappa}, \quad (4.26)$$

where $L = \sup_i |M_i - L_i| \vee 1$.

Proof. We fix p and write I as a shorthand for $I_{2p}(f)$. The properties of cumulants show that, setting $X_i = f_i(Y_{t_i})$ as previously, I is given by

$$I = \sum_{\Delta \in \mathcal{P}([2p])} I_\Delta,$$

$$I_\Delta \stackrel{\text{def}}{=} \int_{L_1}^{M_1} \cdots \int_{L_{2p}}^{M_{2p}} \left(\prod_{A \in \Delta} \mathbf{E}_c X_A \right) \left(\prod_{k=1}^{p} |t_{2k} - t_{2k-1}|^{2H-2} \right) dt_1 \cdots dt_{2p}.$$

Note first that since the f_i are centred, we have $I_\Delta = 0$ unless $|A| \geq 2$ for every $A \in \Delta$. There is furthermore one special partition, namely $\Delta_\star = \{\{2k-1, 2k\} : k \in [p]\}$. For the summand generated by this 'base' partition we have $I_{\Delta_\star} = \prod_{k=1}^{p} I(2k-1, 2k)$, where we set

$$I(k, \ell) = \int_{L_k}^{M_k} \int_{L_\ell}^{M_\ell} \mathbf{E}\big(f_k(Y_s) f_\ell(Y_t)\big) |t - s|^{2H-2} dt \, ds.$$

We then note that, for $a < b$ and $f, g \in E$ centred, it follows from the spectral gap assumption and the fact that $E, E_1 \subset L^2(\mu)$ by Assumption 2.3, that

$$\left| \int_a^b \mathbf{E}\big(f(Y_0)g(Y_t)\big) |t|^{2H-2} dt - C(f, g) \mathbf{1}_{0 \in [a,b]} \right| \lesssim \|f\|_E \|g\|_E e^{-c(|a| \wedge |b|)}, \quad (4.27)$$

for some fixed constant c. It follows from (4.27) that

$$\left| I(k, \ell) - |[L_k, M_k] \cap [L_\ell, M_\ell]| C(f_k, f_\ell) \right| \lesssim \int_{L_k}^{M_k} e^{-c(|L_\ell - s| \wedge |M_\ell - s|)} ds \lesssim 1,$$

and that $I_{\Delta_\star}$ differs from the desired expression in the statement by an error of at most $\mathcal{O}(L^{p-1})$.

Since $I = \sum_{\Delta \in \mathcal{P}([2p])} I_\Delta$ and we already obtained (4.26) for I replaced by $I_{\Delta_\star}$, it remains to show that $|I_\Delta| \lesssim L^{p-\kappa}$ for every partition $\Delta \neq \Delta_\star$ with κ as in the statement. Fix such a partition Δ from now on and write again $\sim$ for the equivalence relation induced by Δ on $[2p]$. We then define a graph $\mathcal{G}_\Delta$ with vertex set $\mathcal{V} = [2p]$ and edge set $\mathcal{E} = \mathcal{E}_B \cup \mathcal{E}_\Delta$, where

$$\mathcal{E}_B = \{(2k-1, 2k) : k \in [p]\}, \qquad \mathcal{E}_\Delta = \{(u, v) : u \sim v\}.$$

We furthermore assign kernels to the edges of $\mathcal{G}_\Delta$ by

$$K_e(t) = \begin{cases} |t|^{2H-2} & \text{for } e \in \mathcal{E}_B, \\ e^{-c|t|} & \text{otherwise,} \end{cases} \quad (4.28)$$

so that Proposition 4.8 yields the bound

$$|I_\Delta| \lesssim \int_{[0,L]^{2p}} \left|K^{\mathcal{G}_\Delta}(t)\right| dt.$$

The kernel assignment (4.28) is consistent with the exponents given by

$$\alpha_-(e) = \begin{cases} 2H - 2 & \text{for } e \in \mathcal{E}_B, \\ 0 & \text{otherwise,} \end{cases} \qquad \alpha_+(e) = \begin{cases} 2H - 2 & \text{for } e \in \mathcal{E}_B, \\ -2 & \text{otherwise,} \end{cases}$$

whence it immediately follows that $\mathcal{G}_\Delta$ is regular.

It now remains to find a function $\beta \colon \mathcal{E} \to \mathbf{R}_+$ allowing us to apply Corollary 4.13 to $\mathcal{G}_\Delta$. For this, we construct a set $\mathcal{T}$ and set $\beta(e) = 1 - \kappa$ for $e \in \mathcal{T}$ and 0 otherwise. To construct $\mathcal{T}$m consider the graph $\hat{\mathcal{G}}_\Delta$ which has $\hat{\mathcal{V}} := \Delta$ as its vertex set and such that its edge set $\hat{\mathcal{E}}$ is given by

$$\hat{\mathcal{E}} = \{(\pi_\Delta(2k - 1), \pi_\Delta(2k)) \,:\, \pi_\Delta(2k - 1) \neq \pi_\Delta(2k)\},$$

where $\pi_\Delta \colon [2p] \to \Delta$ maps an element to the unique element of the partition Δ that contains it. In other words, $\hat{\mathcal{G}}_\Delta$ is obtained by quotienting $\mathcal{G}_\Delta$ by the partition Δ and then removing self-loops. Let now $\mathcal{T} \subset \mathcal{E}_B$ be such that $\hat{\mathcal{T}} = \pi_\Delta \mathcal{T}$ is a maximal spanning forest for $\hat{\mathcal{G}}_\Delta$. In the case of (4.19) for example, one could take $\mathcal{T} = \{(1, 2), (5, 6)\}$. With κ as in the statement of the proposition, we now set $\beta(e) = 1 - \kappa$ for $e \in \mathcal{T}$ and 0 otherwise. The reason why this choice of β satisfies (4.25) for the graph $\mathcal{G}_\Delta$ is that by construction the labelling $\gamma = \alpha_+ - \beta$ is such that $\mathcal{G}_\Delta$ contains a spanning forest $\tilde{\mathcal{T}}$ consisting of edges e with $\gamma(e) = 2H - 3 + \kappa < -1$. (To build a reduced set of edges from $\mathcal{E} = \mathcal{E}_B \cup \mathcal{E}_\Delta$, we start with $\mathcal{T}$ and then connect its components using edges in $\mathcal{E}_\Delta$.) It then remains to first apply Lemma 4.14 to reduce ourselves to considering $\mathcal{G}_\Delta$ and then apply Lemma 4.15.

Denote now by m the number of connected components of $\hat{\mathcal{G}}_\Delta$ and note that since every element of Δ is of size at least 2, $\hat{\mathcal{G}}_\Delta$ has at most p vertices. It follows that the number of elements in $\mathcal{T}$ equals at most $p - m$, so that Corollary 4.13 yields the bound $I_\Delta \lesssim L^m L^{(1-\kappa)(p-m)} = L^{p-\kappa(p-m)}$, which is bounded by $L^{p-\kappa}$ unless $m = p$. Since the only partition Δ yielding $m = p$ is the complete pairing $\Delta_\star$, the claim follows at once. $\qquad\square$

Corollary 4.17. *Let the assumptions of Proposition 4.16 hold. For $H \in (\frac{1}{2}, 1)$ and $f \in E$ with $\int f(y)\,\mu(dy) = 0$, one has for every $p \geq 1$ the bounds*

$$\|J^\varepsilon_{s,t}(f)\|_{L^{2p}} \lesssim t^{\frac{1}{2}}, \quad \|\mathbb{J}^\varepsilon_{s,t}(f)\|_{L^p} \lesssim t,$$

uniformly over $t \leq T$ (for any fixed $T \geq 1$) and $\varepsilon \leq 1$.

Proof. For integer $p \geq 1$, we note that as a consequence of Wick's theorem and the fact that Y is independent of B, one has the identity,

$$\mathbf{E}\big(J^\varepsilon_{0,t}(f)\big)^{2p} = C_p \varepsilon^p \int_{[0,t/\varepsilon]^{2p}} \mathbf{E}\big(f(Y_{s_1}) \cdots f(Y_{s_{2p}})\big) \prod_{k=1}^p \eta''(s_{2k} - s_{2k-1})\, ds.$$

$$\tag{4.29}$$

where $C_p = 2^{-p}(2p-1)!!$. For $t \geq \varepsilon$ we apply Proposition 4.16 with $L = t/\varepsilon$, so that

$$\left| \mathbf{E}\big(J_{0,t}^\varepsilon(f)\big)^{2p} \right| \lesssim \varepsilon^p (t/\varepsilon)^{p-\kappa} + \varepsilon^p (t/\varepsilon)^p \lesssim \varepsilon^\kappa t^{p-\kappa} + t^p \lesssim t^p.$$

For $t < \varepsilon$, (4.29) is bounded by $C\varepsilon^p \|f\|_{L^{2p}}^{2p} |t/\varepsilon|^{2Hp} \lesssim t^p$. One similarly obtains the bound $\mathbf{E}\big(\mathbb{J}_{0,t}^\varepsilon(f,g)\big)^p \lesssim t^p$, completing the proof. □

5. Identification of the Limit

In this section we complete the proof of Theorem 3.14 by identifying the limit (in law) of $\mathbf{Z}^\varepsilon$ as $\varepsilon \to 0$. The proof proceeds in two steps. First, in Proposition 5.3, we show that the first-order process Z itself converges in law to a limit W with covariance given as in (1.8). In a second step, we then exploit martingale techniques, and in particular [39], to obtain convergence of the second-order process $\mathbb{Z}$ to the limit described in Theorem 3.14. Recall that, by (3.17),

$$\big(Z_{s,t}^\varepsilon(x)\big)_i = \varepsilon^{\frac{1}{2}-H} Z_{s,t}^{f_{i,x}^\varepsilon}, \quad \big(\mathbb{Z}_{s,t}^\varepsilon(x,\bar{x})\big)_{ij} \overset{\text{def}}{=} \varepsilon^{1-2H} \mathbb{Z}_{s,t}^{f_{i,x}^\varepsilon, f_{j,\bar{x}}^\varepsilon}.$$

Proposition 5.1. *In the setting of Proposition 4.1, the family of random rough paths $\mathbf{Z}^\varepsilon$ converges in law, as $\varepsilon \to 0$, to the unique (in law) random rough path $\mathbf{Z}$ such that the following hold. The process Z is a $\mathcal{B}$-valued Wiener process with covariance given by*

$$\mathbf{E}\big(Z_{s,t}(x) \otimes Z_{u,v}(\bar{x})\big) = |[s,t] \cap [u,v]| \big(\Sigma(x,\bar{x}) + \Sigma(\bar{x},x)^\top\big), \tag{5.1}$$

with Σ as defined in (1.6). The "second-order" process $\mathbb{Z}$ is the $\mathcal{B}_2$-valued process such that for any $x, \bar{x}$ in $\mathbf{R}^d$

$$\mathbb{Z}_{s,t}(x,\bar{x}) = \int_s^t Z_{s,r}(x) \otimes dZ_{s,r}(\bar{x}) + (t-s)\Sigma(x,\bar{x}),$$

where the integral is interpreted in the Itô sense.

Proof. The convergence in distribution of any finite collections of the stochastic processes follows from Proposition 5.11 below. By Proposition 4.1, $(Z^\varepsilon, \mathbb{Z}^\varepsilon)$ is tight in $\mathscr{C}^\alpha([0,T], \mathcal{B} \oplus \mathcal{B}_2)$ for suitable $\alpha \in (\frac{1}{3}, H)$, so the weak convergence holds with respect to the rough path norm on $\mathscr{C}^\alpha([0,T], \mathcal{B} \oplus \mathcal{B}_2)$. □

5.1. Law of large numbers. We will need the following quantitative version of the law of large numbers. Let $E \subset E_1 \subset E_2$ be Banach spaces of functions $\mathcal{Y} \to \mathbf{R}$ containing constants and such that pointwise multiplication from $E \times E_1$ into E_2 is continuous.

Lemma 5.2. *Let $E \subset L^4$ and $E_2 \subset L^2$, let the spectral gap condition (2.1) hold for $n = 1, 2$, and let $f, g \in E$. Then, the bound*

$$\left\| \int_0^T f(Y_s) g(Y_{s+t}) \, ds - T\langle f, P_t g\rangle_\mu \right\|_{L^2} \lesssim \sqrt{(1+t)T} \|f\|_E \|g\|_E, \tag{5.2}$$

holds uniformly over $t, T \in \mathbf{R}_+$ with a proportionality constant depending only on the constants appearing in the two assumptions.

Proof. Writing f_s as a shorthand for $f(Y_s)$ and similarly for g, the square of the left-hand side of (5.2) is given by

$$2 \int_0^T \int_0^r \big(\mathbf{E}(f_s g_{s+t} f_r g_{r+t}) - \mathbf{E}(f_s g_{s+t}) \mathbf{E}(f_r g_{r+t}) \big) \, ds \, dr.$$

Since $E \subset L^4$, Hölder's inequality shows that the integrand is bounded by some multiple of $\|f\|_E^2 \|g\|_E^2$. It thus follows from the triangle inequality that the required bound follows for $t \geq T$ so we assume $t \leq T$ from now on. Using the same bound on the integrand, we can further restrict the inner integral to impose $s + t \leq r$ at an additive cost of order at most $tT \|f\|_E^2 \|g\|_E^2$. On that smaller domain, we can then rewrite the integrand as

$$\mathbf{E}\big(f_s g_{s+t} (P_{r-s-t}(f P_t g) - \langle f, P_t g \rangle_\mu)(Y_{s+t}) \big).$$

By the spectral gap assumption applied to $f P_t g \in E_2$, we have the bound

$$\|P_{r-s-t}(f P_t g) - \langle f, P_t g \rangle_\mu\|_{E_2} \lesssim e^{-c(r-s-t)} \|f P_t g\|_{E_2} \lesssim e^{-c(r-s-t)} \|f\|_E \|g\|_{E_1}.$$

Combining this again with Hölder's inequality, $E \subset L^4$, and $E_2 \subset L^2$, we conclude that the integrand is of order $e^{-c(r-s-t)} \|f\|_E^2 \|g\|_E^2$, thus yielding a contribution to the integral of order $T \|f\|_E^2 \|g\|_E^2$ as desired. □

5.2. Identification of the first-order process .

We treat separately the cases $H < \frac{1}{2}$ and $H > \frac{1}{2}$, while the case $H = \frac{1}{2}$ is straightforward and will be considered when we put both cases together in Proposition 5.11 below.

5.2.1. The low regularity case

Let $H \in (\frac{1}{3}, \frac{1}{2})$. Conditional on Y, the process $J_{s,t}^\varepsilon(f) = \varepsilon^{\frac{1}{2}-H} \int_s^t f(Y_r^\varepsilon) dB(r)$ is centred Gaussian. In order to identify its limiting distribution, it thus suffices to show that its conditional covariances converge to a limit that is independent of Y. This is the content of the following result.

Proposition 5.3. *Let Assumptions* 2.1 *and* 2.3 *hold for* $H \in (\frac{1}{3}, \frac{1}{2})$, *and let Assumptions* 2.2 *hold for some* $p_\star \geq 4$. *Let* $f, g \in E$ *and let* $u < v$ *and* $s < t$. *Then, we have*

$$\lim_{\varepsilon \to 0} \mathbf{E}\big(J_{s,t}^\varepsilon(f) J_{u,v}^\varepsilon(g) \,|\, \mathcal{F}^Y \big) = |[s, t] \cap [u, v]| \, C(f, g),$$

in L^2, *where* $C(f, g) = \frac{1}{2} \Gamma(2H + 1)\big(\langle f, \mathcal{L}^{1-2H} g \rangle_\mu + \langle \mathcal{L}^{1-2H} f, g \rangle_\mu \big)$.

Proof. We work in components it is therefore again sufficient to assume that $m = 1$. We first consider the case $[u, v] = [s, t]$. A straightforward calculation similar to the one given in (3.13) shows that, setting $\alpha_H = H(1 - 2H)$, one has the identity

$$\mathbf{E}\big(J_{s,t}^\varepsilon(f) J_{s,t}^\varepsilon(g) \,|\, \mathcal{F}^Y \big)$$

$$= \frac{-\alpha_H}{2} \varepsilon^{1-2H} \int_{s-t}^{t-s} |v|^{2H-2} \int_{2s+|v|}^{2t-|v|} \big(f(Y_{\frac{u+v}{2}}^\varepsilon) g(Y_{\frac{u-v}{2}}^\varepsilon) - f(Y_{\frac{u}{2}}^\varepsilon) g(Y_{\frac{u}{2}}^\varepsilon) \big) \, du \, dv$$

$$+ H \varepsilon^{1-2H} \int_0^{t-s} u^{2H-1} \big((fg)(Y_{(2t-u)/2}^\varepsilon) + (fg)(Y_{(2s+u)/2}^\varepsilon) \big) \, du.$$

$$(5.3)$$

Provided that $2p \le p_\star$, the size of the L^p-norm of the last term is at most of order $\|f\|_E\|g\|_E|t - s|^{2H}\varepsilon^{1-2H}$, so that it does not contribute to the limit considered in the statement. (Note however that in the special case $H = \frac{1}{2}$ which we do not consider here, this term is the only surviving one, which contributes to the fact that the conclusion of the lemma still holds in this case.)

Regarding the first term, we note that changing the sign of v is the same as swapping f and g. Taking $s = 0$ (by stationarity) and performing a change of variables, it remains to show that, for any fixed $t > 0$,

$$\varepsilon \int_0^{\frac{t}{\varepsilon}} \int_v^{\frac{2t}{\varepsilon}-v} v^{2H-2}\big(f(Y_{(u+v)/2})g(Y_{(u-v)/2}) - f(Y_{\frac{u}{2}})g(Y_{\frac{u}{2}})\big)\, du\, dv$$

$$\to -t\frac{\Gamma(2H+1)}{H(1-2H)}\langle f, \mathcal{L}^{1-2H}g\rangle_\mu \tag{5.4}$$

in L^2 as $\varepsilon \to 0$. We set

$$\hat{I}_\varepsilon = -\frac{1}{2}\alpha_H\varepsilon\int_0^{\frac{t}{\varepsilon}} v^{2H-2}G_v\, dv,$$

with

$$G_v = \int_v^{\frac{2t}{\varepsilon}-v}\big(f(Y_{(u+v)/2})g(Y_{(u-v)/2}) - f(Y_{\frac{u}{2}})g(Y_{\frac{u}{2}})\big)\, du.$$

To show that $\lim_{\varepsilon\to 0}\hat{I}_\varepsilon = \frac{t}{2}\Gamma(2H+1)\langle f, \mathcal{L}^{1-2H}g\rangle_\mu$ in L^2, we treat the values of v close to the singularity separately from the others, so we fix some (eventually sufficiently small) exponent κ. For "small" values of v, we then have the bound

$$\varepsilon\left\|\int_0^{\varepsilon^\kappa} v^{2H-2}G_v\, dv\right\|_{L^2} \lesssim t\|f\|_E\|g\|_E\int_0^{\varepsilon^\kappa} v^{2H-2}\big(1 \wedge v^H\big)\, dv$$

$$\lesssim \varepsilon^{(3H-1)\kappa}t\|f\|_E\|g\|_E,$$

which converges to 0 as desired for any fixed $\kappa > 0$ since $H > \frac{1}{3}$.

For the remaining values of v, we apply Lemma 5.2, which yields the bound

$$\left\|G_v - 2\Big(\frac{t}{\varepsilon} - |v|\Big)\big(\langle P_v f, g\rangle_\mu - \langle f, g\rangle_\mu\big)\right\|_{L^2} \lesssim \sqrt{\varepsilon^{-1}(1+|v|)(t-\varepsilon|v|)}\|f\|_E\|g\|_E.$$

For $\kappa < \frac{1}{2(1-2H)}$, we furthermore have the bound

$$\varepsilon\int_{\varepsilon^\kappa}^{\frac{t}{\varepsilon}} v^{2H-2}\sqrt{\varepsilon^{-1}(1+|v|)(t-\varepsilon|v|)}\, dv \lesssim \sqrt{\varepsilon t}\int_{\varepsilon^\kappa}^{\frac{t}{\varepsilon}}\big(v^{2H-2} + v^{2H-\frac{3}{2}}\big)\, dv$$

$$\lesssim \sqrt{\varepsilon t}\Big(\int_{\varepsilon^\kappa}^{\infty} v^{2H-2}\, dv + \int_0^{\frac{t}{\varepsilon}} v^{2H-\frac{3}{2}}\, dv\Big) \lesssim \varepsilon^{\frac{1}{2}-\kappa(1-2H)}\sqrt{t} + \varepsilon^{1-2H}t^{2H},$$

which converges to 0 as $\varepsilon \to 0$ for every fixed t. We conclude that

$$
\begin{aligned}
\lim_{\varepsilon \to 0} \hat{I}_\varepsilon &= -\alpha_H \lim_{\varepsilon \to 0} \int_{\varepsilon^\kappa}^{\frac{t}{\varepsilon}} v^{2H-2}\big(t - \varepsilon|v|\big)\big(\langle P_v f, g \rangle_\mu - \langle f, g \rangle_\mu\big)\, dv \\
&= -\alpha_H \lim_{\varepsilon \to 0} \int_0^\infty v^{2H-2} t\big(\langle P_v f, g \rangle_\mu - \langle f, g \rangle_\mu\big)\, dv \\
&= \frac{t}{2}\Gamma(2H+1)\langle f, \mathcal{L}^{1-2H} g \rangle_\mu,
\end{aligned}
$$

holds in L^2, as claimed.

We now consider the case when $[s, t) \cap [u, v) = \emptyset$ and assume without loss of generality that $t \le u$. We then have

$$
\begin{aligned}
\big|\mathbf{E}\big(I_{s,t}^\varepsilon(f) I_{u,v}^\varepsilon(g) \,|\, \mathcal{F}^Y\big)\big| &= \alpha_H \left| \int_s^t \int_u^v |r - \bar{r}|^{2H-2} f(Y_r^\varepsilon) g(Y_{\bar{r}}^\varepsilon)\, d\bar{r}\, dr \right| \\
&\lesssim \int_s^t \int_u^v |r - \bar{r}|^{2H-2}\, d\bar{r}\, dr < \infty.
\end{aligned}
$$

Since this is multiplied by ε^{1-2H}, it follows that $\mathbf{E}\big(J_{s,t}^\varepsilon(f) J_{u,v}^\varepsilon(g) \,|\, \mathcal{F}^Y\big) \to 0$ in that case. The general case then follows immediately since we have $J_{s,t}^\varepsilon(f) = J_{s,u}^\varepsilon(f) + J_{u,t}^\varepsilon(f)$ for any $s \le u \le t$, so that it can be reduced to the two cases we just treated. $\quad\square$

5.2.2. The high regularity case Let $H > \frac{1}{2}$. We first show that the process Z is Gaussian with covariance given by (5.1). For this we recall relations between cumulants and expectations. Fix a finite set A as well as elements $f_a \in E$ and intervals $[s_a, t_a] \subset \mathbf{R}$ for every $a \in A$. Given a subset $B \subset A$, we write $\mathcal{G}(B)$ for the set of pairs (Δ, p) where $\Delta \in \mathcal{P}(B)$ is a partition of B without singletons and p is a pairing of B (i.e. $p \in \mathcal{P}(B)$ contains only sets of size two). We also write $[s, t]_B \subset \mathbf{R}^B$ for the domain $\times_{a \in B}[s_a, t_a]$. Given $G = (\Delta, p) \in \mathcal{G}(B)$, we then set

$$
J_G^\varepsilon = (-\varepsilon^{\frac{1}{2}-H}\alpha_H)^{|B|} \int_{[s,t]_B} \prod_{B' \in \Delta} \mathbf{E}_c\big(f_a(Y_{r_a}^\varepsilon) : a \in B'\big) \prod_{\{a,b\} \in p} |r_a - r_b|^{2H-2}\, dr.
$$

In order to extract a formula for the joint cumulants of the $J_{s,t}^\varepsilon$'s, we note that if $B_1 \cap B_2 = \emptyset$ and $G_i \in \mathcal{G}(B_i)$, one has

$$
J_{G_1 \sqcup G_2}^\varepsilon = J_{G_1}^\varepsilon \cdot J_{G_2}^\varepsilon, \tag{5.5}
$$

where $G_1 \sqcup G_2 \in \mathcal{G}(B_1 \sqcup B_2)$ denotes the natural concatenation of G_1 and G_2. We furthermore write $\mathcal{G}_c(B) \subset \mathcal{G}(B)$ for the set of "connected" elements, namely those pairs $G = (\Delta, p)$ such that $G^\vee \stackrel{\text{def}}{=} \Delta \vee p = \{B\}$, where we use the usual lattice structure of the set of partitions of B.

Lemma 5.4. *Let Assumptions* 2.3 *holds for* $H > \frac{1}{2}$ *and* $f_a \in E$, *then for every* $B \subset A$,

$$
\mathbf{E}_c\big(J_{s_a,t_a}^\varepsilon(f_a) : a \in B\big) = \sum_{G \in \mathcal{G}_c(B)} J_G^\varepsilon.
$$

Proof. As a consequence of (5.5), we can write

$$\mathbf{E}\prod_{a\in A} J^{\varepsilon}_{s_a,t_a}(f_a) = \sum_{G\in\mathcal{G}(A)} J^{\varepsilon}_G = \sum_{\Delta'\in\mathcal{P}(A)}\sum_{G\in\mathcal{G}(A):G^{\vee}=\Delta'} J^{\varepsilon}_G$$

$$= \sum_{\Delta'\in\mathcal{P}(A)}\prod_{B\in\Delta'}\sum_{G\in\mathcal{G}_c(B)} J^{\varepsilon}_G.$$

Comparing this to the first identity in (4.15), we conclude the proof. □

This allows to conclude:

Lemma 5.5. *Let Assumptions* 2.1 *and* 2.3 *hold for* $H > \frac{1}{2}$. *Let* $f_a \in E$ *with* $\int f_a d\mu = 0$. *Then, for any index set* B *with more than two elements,*

$$\mathbf{E}_c\big(\lim_{\varepsilon\to 0} J^{\varepsilon}_{s_a,t_a}(f_a) : a\in B\big) = 0. \qquad (5.6)$$

Consequently, the limits $\{J_{s_a,t_a}(f_a)\}_{a\in B}$ *are jointly Gaussian with covariance* $|[s_a, t_a]\cap [s_b, t_b]|\, C(f_a, f_b)$.

Proof. Let $G = (\Delta, p) \in \mathcal{G}_c(B)$ with $|B|$ even and note that J^{ε}_G can be written as

$$J^{\varepsilon}_G = (-\alpha_H)^{|B|}\varepsilon^{|B|/2}\int_{[\frac{s}{\varepsilon},\frac{t}{\varepsilon}]^B}\prod_{\bar{A}\in\Delta}\mathbf{E}_c\big(f_a(Y_{r_a}) : a\in\bar{A}\big)\prod_{\{a,b\}\in p}|r_a - r_b|^{2H-2}\, dr.$$

We then note that if $|B| > 2$, elements $(\Delta, p) \in \mathcal{G}_c(B)$ are always such that $\Delta \neq p$. We can therefore apply Proposition 4.16 with $2p = |B|$, which shows that

$$|J^{\varepsilon}_G| \lesssim \varepsilon^{|B|/2}\varepsilon^{\kappa-\frac{|B|}{2}},$$

which converges to 0, thus yielding (5.6) as claimed. By Proposition 4.16, for any $f_i \in E$, $u < v$ and $s < t$ we have

$$\lim_{\varepsilon\to 0}\Big|\mathbf{E}\big(J^{\varepsilon}_{s,t}(f_i)J^{\varepsilon}_{u,v}(f_j)\big) - |[s, t]\cap[u, v]|\, C(f_i, f_j)\Big| = 0.$$

Since Gaussian processes are characterised by the fact that their joint cumulants of order three or higher all vanish, the last claim follows. □

5.3. Convergence of the second-order process .

In this section we assume that $H \in (\frac{1}{3}, 1)$. If $H = \frac{1}{2}$, $J^{\varepsilon}_t = \int_0^t f(Y^{\varepsilon}_r)dB_r$ is already a local martingale, otherwise we make a decomposition as follows. Write $\bar{B}^t_r = \mathbf{E}\big(B_r - B_t \mid \mathcal{F}^B_t\big)$ for $r > t$ and write

$$J^{\varepsilon}_t(f) = \varepsilon^{\frac{1}{2}-H}\int_0^t f(Y^{\varepsilon}_r)dB_r = M^f_t + R^f_t,$$

where, setting $\bar{f} = \int_{\mathcal{Y}} f(x)\,\mu(dx)$ and $\tilde{f} = f - \bar{f}$,

$$M^f_t = \varepsilon^{\frac{1}{2}-H}\int_0^t \tilde{f}(Y^{\varepsilon}_r)\, dB_r + \varepsilon^{\frac{1}{2}-H}\int_t^{\infty}\big(P_{\frac{r-t}{\varepsilon}}\tilde{f}\big)(Y^{\varepsilon}_t)\, d\bar{B}^t_r, \qquad (5.7)$$

$$R^f_t = \varepsilon^{\frac{1}{2}-H}\bar{f}B_t - \varepsilon^{\frac{1}{2}-H}\int_t^{\infty}\big(P_{\frac{r-t}{\varepsilon}}\tilde{f}\big)(Y^{\varepsilon}_t)\, d\bar{B}^t_r. \qquad (5.8)$$

The convergence of the integral is guaranteed by the fact that $|\dot{\bar{B}}_t^r| \lesssim (r-t)^{(H-1)-}$ and $P_t \tilde{f} \to 0$ exponentially fast, thanks to the centering condition on $\tilde{f}$. This clearly illustrates why have no need to assume that the coefficient $F(x, \cdot)$ itself is centred for $H < \frac{1}{2}$ since the first term in R_t^f obviously converges to 0 in that case. For $H = \frac{1}{2}$, we write $M_t^f = J_t^\varepsilon(f)$ and $R_t^f = 0$.

Lemma 5.6. *Let $H \in (\frac{1}{3}, 1)$, let $p > 1$ and $E \subset L^{2p}$, and let* (2.1) *holds for $n = 1$ For every $f \in E$ and and every $t \geq 0$, one has $\lim_{\varepsilon \to 0} R_t^f = 0$ in L^q for any $q \in [1, p)$. Furthermore, $\left\| \varepsilon^{-H} \int_t^\infty (P_{\frac{r-t}{\varepsilon}} \tilde{f})(Y_t^\varepsilon) d\bar{B}_r^t \right\|_{L^q} \lesssim \|\tilde{f}\|_E$.*

Proof. We only need to prove if for $H \neq \frac{1}{2}$. The first term in the definition of R_t^f obviously converges to 0. The scale and shift invariance of fractional Brownian motion shows that, in law, the second term $\varepsilon^{\frac{1}{2}-H} \int_t^\infty (P_{\frac{r-t}{\varepsilon}} \tilde{f})(Y_t^\varepsilon) d\bar{B}_r^t$ equals

$$\varepsilon^{\frac{1}{2}} \int_0^\infty (P_r \tilde{f})(Y) d\bar{B}_r^0,$$

with Y a random variable with law μ, independent of $\bar{B}^0$. Note now that $\dot{\bar{B}}^0$ is Gaussian and

$$\mathbf{E}|\dot{\bar{B}}_r^0|^2 \propto |r|^{2H-2}.$$

Furthermore since $f \in E$, $\|P_t f\|_{L^p} \lesssim |f|_E$ and $\|P_t f\|_{L^1} \to 0$, $\|P_t f\|_{L^q} \to 0$ for any $q \in [1, p)$, so that by Cauchy–Schwarz,

$$\left\| \int_0^\infty (P_r \tilde{f})(Y) d\bar{B}_r^0 \right\|_{L^p} \lesssim \int_0^\infty \|(P_r \tilde{f})(Y)\|_{L^{2p}} |r|^{H-1} dr$$

$$\lesssim \|\tilde{f}\|_E \int_0^\infty e^{-cr} |r|^{H-1} dr < \infty,$$

and the claim follows. $\square$

Lemma 5.7. *Let $H \in (\frac{1}{3}, 1) \backslash \{\frac{1}{2}\}$. Let Assumption* 2.1 *hold for $n = 1$, let Assumption* 2.2 *hold for some $p_\star > 2$, and let Assumption* 2.3 *hold.*

Let $f \in E$, then the process M_t^f as defined above is an L^p bounded $\mathcal{F}_t^Y \vee \mathcal{F}_t^B$-martingale for every $p < p_\star$ with the convention that $p_\star = \infty$ for $H > \frac{1}{2}$).

Proof. For $T > 0$, we define the $\mathcal{F}_t^Y \vee \mathcal{F}_t^B$-martingale $M_t^{f,T}$ by

$$M_t^{f,T} = \varepsilon^{\frac{1}{2}-H} \mathbf{E}\left(\int_0^T \tilde{f}(Y_r^\varepsilon) dB_r \,\Big|\, \mathcal{F}_t^Y \vee \mathcal{F}_t^B \right),$$

and we note that for $T > t$ one has

$$M_t^{f,T} = \varepsilon^{\frac{1}{2}-H} \int_0^t \tilde{f}(Y_r^\varepsilon) dB_r + \varepsilon^{\frac{1}{2}-H} \int_t^T (P_{\frac{r-t}{\varepsilon}} \tilde{f})(Y_t^\varepsilon) d\bar{B}_r^t.$$

Since $\|P_t \tilde{f}\|_{E_1} \to 0$ exponentially fast, it follows from Lemma 5.6 that $M_t^f = \lim_{T \to \infty} M_t^{f,T}$ in L^p, so that M_t^f is a local martingale. Since the first term of (5.7) is bounded in L^p (by Proposition 4.5 for $H \in (\frac{1}{3}, \frac{1}{2})$, and for $H > \frac{1}{2}$ from Corollary 4.17 which applies since $\tilde{f}$ is centred), and the second term converges in L^p by Lemma 5.6, the claim follows. $\square$

Remark 5.8. For $H < \frac{1}{2}$, we only used the integrability condition $E \subset L^{p_\star}$ in Lemma 5.7, the condition $E_2 \subset L^2$ from Assumption 2.3 is not needed (nor is it needed in Proposition 4.5).

For $H \neq \frac{1}{2}$, we can then rewrite $\mathbb{J}$ as

$$\mathbb{J}_{0,t}^\varepsilon(f, g) = \int_0^t J_s^\varepsilon(f) \, dM_s^g + J_t^\varepsilon(f) R_t^g - \int_0^t R_s^g \, dJ_s^\varepsilon(f). \tag{5.9}$$

Here, the first integral is an Itô integral, while the integral $\int_0^t J_s^\varepsilon(f) dR_s^f$ should be interpreted as the limit, as $\delta \to 0$, of the corresponding expression with B replaced by a mollified version B^δ, to which integration by parts is applied.

Lemma 5.9. *Let $H \in (\frac{1}{3}, \frac{1}{2}) \cup (\frac{1}{2}, 1)$. Let Assumptions 2.1–2.3 hold and let $f \in E$. Then, one has*

$$\lim_{\varepsilon \to 0} \mathbf{E}\left(\int_0^t R_s^g \, dJ_s^\varepsilon(f) \,\Big|\, \mathcal{F}^Y \right) = -\frac{1}{2} t \Gamma(2H + 1) \langle \mathcal{L}^{1-2H} g, f \rangle_\mu,$$

in probability.

Proof. Let us first write $R_t^g = J_t^\varepsilon(\bar{g}) - \tilde{R}_t^g$ where

$$\tilde{R}_t^g = \varepsilon^{\frac{1}{2}-H} \int_t^\infty \left(P_{\frac{r-t}{\varepsilon}} \tilde{g} \right)(Y_t^\varepsilon) \, d\bar{B}_r^t,$$

and note that

$$\int_0^t J_s^\varepsilon(\bar{g}) \, dJ_s^\varepsilon(f) = J_t^\varepsilon(\bar{g}) \, J_t^\varepsilon(f) - \mathbb{J}_{0,t}^\varepsilon(f, \bar{g}), \tag{5.10}$$

for $H < \frac{1}{2}$ (vanishes for $H > \frac{1}{2}$ since then $\bar{g} = 0$). Since $J_t^\varepsilon(\bar{g}) = \varepsilon^{\frac{1}{2}-H} \bar{g} B_t$, we conclude from Proposition 4.5 that the first term on the right hand side converges to 0 in probability. Since $\bar{g}$ is constant, the second part of Proposition 4.7 implies that the second term also converges to 0 in probability for $H \in (\frac{1}{3}, \frac{1}{2})$, so that it remains to obtain the limit of $\int_0^t \tilde{R}_s^g \, dJ_s^\varepsilon(f)$. For $H \neq \frac{1}{2}$ we have the identity

$$\mathbf{E}\left(\int_0^t \tilde{R}_s^g \, dJ_s^\varepsilon(f) \,\Big|\, \mathcal{F}^Y \right) = \varepsilon^{1-2H} \mathbf{E}\left(\int_0^t \int_s^\infty \left(P_{\frac{r-s}{\varepsilon}} \tilde{g} \right)(Y_s^\varepsilon) d\bar{B}_r^s f(Y_s^\varepsilon) \, dB_s \,\Big|\, \mathcal{F}^Y \right)$$

$$= \frac{\varepsilon^{1-2H}}{2} \int_0^t \int_s^\infty \left(P_{\frac{r-s}{\varepsilon}} \tilde{g} \right)(Y_s^\varepsilon) f(Y_s^\varepsilon) \, \eta''(s - r) \, dr \, ds$$

$$= H(2H - 1) \int_0^t \int_0^\infty \left(P_r \tilde{g} - \mathbf{1}_{H < \frac{1}{2}} \tilde{g} \right)(Y_s^\varepsilon) \, r^{2H-2} \, dr f(Y_s^\varepsilon) \, ds$$

$$= \frac{1}{2} \Gamma(2H + 1) \int_0^t \langle \mathcal{L}^{1-2H} g, f \rangle_\mu (Y_s^\varepsilon) \, ds.$$

We have used the fact that the difference between $B_r - B_s$ and $\bar{B}_r^s$ is independent of B_s. The claim now follows from Birkhoff's ergodic theorem (or the quantitative version given in Lemma 5.2). $\qquad\square$

Proposition 5.10. *Let* $H \in (\frac{1}{3}, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, *let Assumptions* 2.1–2.3 *hold, and let* $f \in E$. *One has*

$$\lim_{\varepsilon \to 0} \int_0^t R_s^g \, dJ_s^\varepsilon(f) = -\frac{t}{2} \Gamma(2H - 1) \langle \mathcal{L}^{1-2H} g, f \rangle,$$

in probability.

Proof. Since the expectation converges by Lemma 5.9, it remains to show that the variance vanishes as $\varepsilon \to 0$. As in the proof of Proposition 4.7, we can fix a realisation of Y and use Proposition 3.4 to reduce ourselves to the case where, conditional on Y, R_t^g and $J_t^\varepsilon(f)$ are independent. As a consequence of the proof of Lemma 5.9, $\int_0^t J_s^\varepsilon(\bar{g}) \, dJ_s^\varepsilon(f) \to 0$ in probability, so it suffices to bound the conditional variance of the term with R_t^g replaced by $\tilde{R}_t^g$.

Writing $A_t = \int_0^t \tilde{R}_s^g \, dJ_s^\varepsilon(f)$ (and assuming that $\tilde{R}^g$ and $J^\varepsilon(f)$ are driven by independent fractional Brownian motions), we then have as in (4.8) the identity

$$\mathbf{E}\big(A_t^2 \mid \mathcal{F}^Y\big) = \frac{1}{2} \int_0^t \int_0^t \phi_\varepsilon(s, s') \eta''(s - s') \, ds' \, ds$$

but this time we have

$$\phi_\varepsilon(s, s') = \varepsilon^{2-4H} f_{s'}^\varepsilon f_s^\varepsilon \int_s^\infty \int_{s'}^\infty \big(P_{\frac{r'-s'}{\varepsilon}} \tilde{g}\big)_{s'}^\varepsilon \big(P_{\frac{r-s}{\varepsilon}} \tilde{g}\big)_s^\varepsilon C_{s \wedge s'}(r, r') \, dr' \, dr, \quad (5.11)$$

where we use the shorthand $f_s^\varepsilon = f(Y_s^\varepsilon)$ and where

$$C_{s \wedge s'}(r, r') = \mathbf{E}\dot{\tilde{B}}_r^s \dot{\tilde{B}}_{r'}^{s'} = \int_{-\infty}^{s \wedge s'} (r - u)^{H - \frac{3}{2}} (r' - u)^{H - \frac{3}{2}} \, du.$$

Note that this holds for any $(\frac{1}{3}, \frac{1}{2}) \cup (\frac{1}{2}, 1)$ and for $r, r' \geq s$, we have the bound $|C_s(r, r')| \lesssim |r - s|^{H-1} |r' - s|^{H-1}$, which will be used repeatedly below. As a consequence of this, the $s \leftrightarrow s'$ symmetry of the integrand in (5.11) and Assumptions 2.1, we obtain for $p_\star \geq 4$ the upper bound

$$\begin{aligned}
\|\phi_\varepsilon(s, s)\|_{L^1} &\lesssim \varepsilon^{2-4H} \int_s^\infty e^{-c(r-s)/\varepsilon} \int_s^r C_s(r, r') \, dr' \, dr \\
&\lesssim \varepsilon^{2-4H} \int_s^\infty e^{-c(r-s)/\varepsilon} \int_s^r |r' - s|^{H-1} |r - s|^{H-1} \, dr' \, dr \\
&\lesssim \varepsilon^{2-4H} \int_s^\infty e^{-c(r-s)/\varepsilon} |r - s|^{2H-1} \, dr \\
&= \varepsilon^{2-2H} \int_0^\infty e^{-cr} r^{2H-1} \, dr \lesssim \varepsilon^{2-2H}, \quad (5.12)
\end{aligned}$$

with constant proportional to $\|f\|_E^2 \|g\|_E^2$. When $H > \frac{1}{2}$, using the local integrability of η'', this is sufficient to conclude that

$$\|\mathbf{E}\big(A_t^2 \mid \mathcal{F}^Y\big)\|_{L^1} \lesssim \varepsilon^{2-2H} \left| \int_0^t \int_0^t \eta''(s - s') \, ds' \, ds \right| \lesssim \varepsilon^{2-2H} t^{2H},$$

which does indeed converge to 0 as $\varepsilon \to 0$ as desired.

It remains to consider the case $H < \frac{1}{2}$, so we restrict ourselves to this case from now on. Regarding $\delta\phi_\varepsilon(s, s') = \phi_\varepsilon(s, s') - \phi_\varepsilon(s, s)$ for $s' > s$ (the case $s < s'$ is analogous), we write $\delta\phi_\varepsilon = \sum \delta\phi_\varepsilon^{(i)}$ with

$$\delta\phi_\varepsilon^{(1)}(s, s') = \varepsilon^{2-4H} f_{s'}^\varepsilon f_s^\varepsilon \int_s^{s'} \int_{s'}^\infty \left(P_{\frac{r'-s'}{\varepsilon}} \tilde{g}\right)_{s'}^\varepsilon \left(P_{\frac{r-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon C_s(r, r') \, dr' \, dr,$$

$$\delta\phi_\varepsilon^{(2)}(s, s') = -\varepsilon^{2-4H} f_s^\varepsilon f_s^\varepsilon \int_s^{s'} \int_s^{s'} \left(P_{\frac{r'-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon \left(P_{\frac{r-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon C_s(r, r') \, dr' \, dr,$$

$$\delta\phi_\varepsilon^{(3)}(s, s') = -\varepsilon^{2-4H} f_s^\varepsilon f_s^\varepsilon \int_s^{s'} \int_{s'}^\infty \left(P_{\frac{r'-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon \left(P_{\frac{r-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon C_s(r, r') \, dr' \, dr,$$

$$\delta\phi_\varepsilon^{(4)}(s, s') = -\varepsilon^{2-4H} f_s^\varepsilon f_s^\varepsilon \int_{s'}^\infty \int_s^{s'} \left(P_{\frac{r'-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon \left(P_{\frac{r-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon C_s(r, r') \, dr' \, dr,$$

$$\delta\phi_\varepsilon^{(5)}(s, s') = \varepsilon^{2-4H} \left(f_{s'}^\varepsilon - f_s^\varepsilon\right) f_s^\varepsilon \int_{s'}^\infty \int_{s'}^\infty \left(P_{\frac{r'-s'}{\varepsilon}} \tilde{g}\right)_{s'}^\varepsilon \left(P_{\frac{r-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon C_s(r, r') \, dr' \, dr,$$

$$\delta\phi_\varepsilon^{(6)}(s, s') = \varepsilon^{2-4H} f_s^\varepsilon f_s^\varepsilon \int_{s'}^\infty \int_{s'}^\infty \left(P_{\frac{r'-s'}{\varepsilon}} \tilde{g} - P_{\frac{r'-s}{\varepsilon}} \tilde{g}\right)_{s'}^\varepsilon \left(P_{\frac{r-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon C_s(r, r') \, dr' \, dr,$$

$$\delta\phi_\varepsilon^{(7)}(s, s') = \varepsilon^{2-4H} f_s^\varepsilon f_s^\varepsilon \int_{s'}^\infty \int_{s'}^\infty \left(\left(P_{\frac{r'-s}{\varepsilon}} \tilde{g}\right)_{s'}^\varepsilon - \left(P_{\frac{r'-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon\right) \left(P_{\frac{r-s}{\varepsilon}} \tilde{g}\right)_s^\varepsilon C_0(r, r') \, dr' \, dr.$$

We obtain the bound

$$\|\delta\phi_\varepsilon^{(1)}(s, s')\|_{L^1} \lesssim \varepsilon^{2-4H} \int_{s'}^\infty e^{-c(r'-s')/\varepsilon} |r' - s|^{H-1} \int_s^{s'} |r - s|^{H-1} \, dr \, dr'$$

$$\lesssim \varepsilon^{2-4H} |s - s'|^H \int_{s'}^\infty e^{-c(r'-s')/\varepsilon} |r' - s'|^{H-1} \, dr'$$

$$\lesssim \varepsilon^{2-3H} |s - s'|^H,$$

and similarly for $\delta\phi_\varepsilon^{(3)}$ and $\delta\phi_\varepsilon^{(4)}$. Regarding $\delta\phi_\varepsilon^{(2)}$, we obtain

$$\|\delta\phi_\varepsilon^{(2)}(s, s')\|_{L^1} \lesssim \varepsilon^{2-4H} \left(\int_s^{s'} e^{-c(r-s)/\varepsilon} |r - s|^{H-1} \, dr\right)^2 \lesssim \varepsilon^{2-4H} |s - s'|^{2H}.$$

In view of (5.12) and Assumption 2.2, we obtain for $\delta\phi_\varepsilon^{(5)}$ the bound

$$\|\delta\phi_\varepsilon^{(5)}(s, s')\|_{L^1} \lesssim \varepsilon^{2-3H} |s' - s|^H,$$

using $\|f_{s'}^\varepsilon - f_s^\varepsilon\|_{L^p} \lesssim (|s - s'|/\varepsilon)^H$ to obtain the increment in time.

In order to bound $\delta\phi_\varepsilon^{(6)}$, we note that one has the bound

$$\|(P_t \tilde{g} - \tilde{g})(Y_s^\varepsilon)\|_{L^p} = \left\|\mathbf{E}\left(\tilde{g}(Y_{t+s}^\varepsilon) - \tilde{g}(Y_s^\varepsilon) \mid \mathcal{G}_s\right)\right\|_{L^p}$$

$$\leq \left\|\tilde{g}(Y_{t+s}^\varepsilon) - \tilde{g}(Y_s^\varepsilon)\right\|_{L^p} \lesssim \|\tilde{g}\|_E \left(1 \wedge t^H \varepsilon^{-H}\right).$$

As a consequence of Assumption 2.1, we thus obtain the bound

$$\|\delta\phi_\varepsilon^{(6)}(s, s')\|_{L^1} \lesssim \varepsilon^{2-3H} |s' - s|^H,$$

and similarly for $\delta\phi_\varepsilon^{(7)}$. Collecting all of these bounds, we conclude that

$$\|\delta\phi_\varepsilon(s, s')\|_{L^1} \lesssim \varepsilon^{2-3H}|s' - s|^H + \varepsilon^{2-4H}|s - s'|^{2H}.$$

It suffices then to apply Lemma 4.3 to ϕ_ε with $\beta = 0$, $\hat{C} = 0$, and $\zeta \in \{H, 2H\}$ to conclude that $\|\mathbf{E}(A_t^2 \mid \mathcal{F}^Y)\|_{L^1} \lesssim \varepsilon^{2-2H}t^{2H} + \varepsilon^{2-4H}t^{4H}$, which converges to 0 as $\varepsilon \to 0$, thus concluding the proof. □

Collecting all of these results, we conclude that the following holds.

Proposition 5.11. *Let Assumptions 2.1–2.3 hold and let $f_1, \ldots f_N \in E$ for some $N \geq 1$. The processes $\left(J_t^\varepsilon(f_i), \mathbb{J}_{s,t}^\varepsilon(f_i, f_j)\right)_{i,j \leq N}$ converge jointly in distribution to*

$$\left(W_t^{(i)}, \int_s^t W_r^{(i)} \, dW^{(j)}(r) + \frac{1}{2}(t - s)\Gamma(2H + 1)\langle f_i, \mathcal{L}^{1-2H} f_j\rangle_\mu\right)_{i,j \leq N},$$

where the $W^{(i)}$ are Wiener processes with covariance

$$\mathbf{E}W_s^{(i)}W_t^{(j)} = \frac{1}{2}(s \wedge t)\Gamma(2H + 1)\left(\langle\mathcal{L}^{1-2H} f_i, f_j\rangle_\mu + \langle\mathcal{L}^{1-2H} f_j, f_i\rangle_\mu\right). \quad (5.13)$$

Proof. If $H < \frac{1}{2}$, by Proposition 5.3 and Lemma 5.6, $(J_t^\varepsilon(f_i), M_t^\varepsilon(f_i), R_t^{f_i})_{i \leq N}$ all converge jointly to $(W_t^{(i)}, W_t^{(j)}, 0)$. This holds similarly for $H > \frac{1}{2}$, using Lemma 5.5. Since by Proposition 5.10 below, $\int_0^t R_s^{f_i} \, dJ_s^\varepsilon(f_j)$ converges to a deterministic limit, and $\int_0^t J_s^\varepsilon(f_i) dM_t^\varepsilon(f_j) \to \int_0^t W_r^{(i)} \, dW_r^{(j)}$, the desired convergence in distribution follows by combining (5.9) with the standard convergence theorem of stochastic integrals, c.f. [39, Theorem 6.22] and [43, Theorem 2.7].

The case $H = \frac{1}{2}$ is straightforward. Firstly, we see that conditional on $\mathcal{F}^Y$, the $J_{s,t}^\varepsilon(f_i) = \int_s^t f_i(Y_r^\varepsilon) \, dB(r)$ are L^2 bounded martingales with respect to the filtration generated by B. By Lemma 5.2, their covariances $\int_s^t \mathbf{E}(f_i f_j)(Y_r^\varepsilon) \, dr$ converge to $(t - s)\langle f_i, f_j\rangle_\mu$ in L^2. Then,

$$\mathbb{J}_{s,t}^\varepsilon = \int_s^t J_r^\varepsilon(f_i) \circ dJ_r^\varepsilon(f_j) = \int_s^t \int_s^u f(Y_r^\varepsilon)g(Y_u^\varepsilon)dB_r dB_u + \frac{1}{2}\int_s^t (fg)(Y_r^\varepsilon)dr,$$

which converges in L^2. Since $J_{s,t}^\varepsilon(f_i)$ converge, they converge together with their integrals, concluding the convergence of $(J_t^\varepsilon(f_i), \mathbb{J}_{s,t}^\varepsilon(f_i, f_j))_{i,j \leq N}$ for $H = \frac{1}{2}$. □

Declarations

Conflict of interest The authors have no conflict of interest to disclose.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A. A Compactness Criterion

Recall the definitions of $\mathcal{B}$ and $\mathcal{B}_2$ from Sect. 2, fix $\zeta \in (0, 1)$ and $\kappa > 0$, and define $\hat{\mathcal{B}} \subset \mathcal{B}$ as the space of $\mathcal{C}^3$ functions X on $\mathbf{R}^d$ such that

$$\|Z\|_{\hat{\mathcal{B}}} \overset{\text{def}}{=} \sup_{x \in \mathbf{R}^d} \left((1 + |x|)^{\kappa/2}\|Z\|_{\mathcal{C}_x^{3+\zeta}}\right) < \infty, \tag{A.1}$$

where

$$\|Z\|_{\mathcal{C}_x^{3+\zeta}} = \sup_{x' : |x'-x| \leq 1} \sup_{|\ell| \leq 3} \left(|D^\ell Z(x)| + \frac{|D^\ell Z(x) - D^\ell Z(x')|}{|x - x'|^\zeta}\right).$$

Similarly, we define $\hat{\mathcal{B}}_2$ as the space of functions on $\mathbf{R}^d \times \mathbf{R}^d$ such that

$$\|Z\|_{\hat{\mathcal{B}}_2} \overset{\text{def}}{=} \sup_{x,\bar{x} \in \mathbf{R}^d} \left((1 + |x| + |\bar{x}|)^{\kappa/2}\|Z\|_{\mathcal{C}_{(x,\bar{x})}^{3+\zeta}}\right) < \infty.$$

We then have the following.

Lemma A.1. *The embeddings $\hat{\mathcal{B}} \subset \mathcal{B}$ and $\hat{\mathcal{B}}_2 \subset \mathcal{B}_2$ are compact for any $\zeta \in (0, 1)$ and $\kappa > 0$. Furthermore, there exists a constant C such that for any random $\mathcal{C}^3$ functions Z and $\bar{Z}$ one has the bound*

$$\mathbf{E}\|Z\|_{\hat{\mathcal{B}}}^p \leq C \sup_{|x-x'| \leq 1} \sup_{|\ell| \leq 3} (1 + |x|)^{\kappa p} \frac{\mathbf{E}|D^\ell Z(x) - D^\ell Z(x')|^p}{|x - x'|^p},$$

$$\mathbf{E}\|\bar{Z}\|_{\hat{\mathcal{B}}_2}^p \leq C \sup_{\substack{|x-x'| \leq 1 \\ |\bar{x}-\bar{x}'| \leq 1}} \sup_{|k+\ell| \leq 3} (1 + |x| + |\bar{x}|)^{\kappa p} \frac{\mathbf{E}|D_1^k D_2^\ell \bar{Z}(x, \bar{x}) - D_1^k D_2^\ell \bar{Z}(x', \bar{x}')|^p}{|x - x'|^p + |\bar{x} - \bar{x}'|^p},$$

$$\tag{A.2}$$

provided that $p \geq d$, $\zeta < 1 - d/p$, and $\kappa > 4d/p$.

Proof. The compactness statement is a routine modification of Arzelá–Ascoli. Regarding the first bound, it follows from Kolmogorov's continuity criterion [64, Thm 2.1, p. 26] that, writing K for the right-hand-side of (A.2), there is $C > 0$ such that

$$\mathbf{E}\|Z\|_{\mathcal{C}_x^{3+\zeta}}^p \leq C(1 + |x|)^{-\kappa p} K,$$

provided that $0 < \zeta < 1 - d/p$. We then cover $\mathbf{R}^d$ with balls of diameter 1 and note that a norm equivalent to that of $\hat{\mathcal{B}}$ is obtained by restricting the supremum in (A.1) over

the centres of these balls. Since $\kappa > 2d/p$, we can then simply bound the supremum by the sum, yielding

$$\mathbf{E}\|Z\|_{\hat{\mathcal{B}}}^p \leq \sum_x \mathbf{E}\big((1+|x|)^{\kappa p/2}\|Z\|_{\mathcal{C}_x^{3+\zeta}}^p\big) \leq C \sum_x (1+|x|)^{-\kappa p/2} K \leq C' K,$$

as claimed. The second bound is identical, except that d is replaced by $2d$. □

References

1. Bailleul, I., Catellier, R.: Rough flows and homogenization in stochastic turbulence. J. Differ. Equ. **263**(8), 4894–4928 (2017). https://doi.org/10.1016/j.jde.2017.06.006
2. Bourguin, S., Gailus, S., Spiliopoulos, K.: Typical dynamics and fluctuation analysis of slow-fast systems driven by fractional Brownian motion. Stochast. Dyn. **21**(7), 2150030 (2021). https://doi.org/10.1142/S0219493721500301
3. Ben Hariz, S.: Limit theorems for the non-linear functional of stationary Gaussian processes. J. Multi-variate Anal. **80**(2), 191–216 (2002). https://doi.org/10.1006/jmva.2001.1986
4. Birrell, J., Hottovy, S., Volpe, G., Wehr, J.: Small mass limit of a Langevin equation on a manifold. Ann. Henri Poincaré **18**(2), 707–755 (2017). https://doi.org/10.1007/s00023-016-0508-3
5. Breuer, P., Major, P.: Central limit theorems for nonlinear functionals of Gaussian fields. J. Multivariate Anal. **13**(3), 425–441 (1983). https://doi.org/10.1016/0047-259X(83)90019-2
6. Bai, S., Taqqu, M.S.: Multivariate limit theorems in the context of long-range dependence. J. Time Ser. Anal. **34**(6), 717–743 (2013). https://doi.org/10.1111/jtsa.12046
7. Chevyrev, I., Friz, P.K., Korepanov, A., Melbourne, I., Zhang, H.: Multiscale systems, homogenization, and rough paths. In: Probability and Analysis in Interacting Physical Systems (2019)
8. Cogburn, R., Hersh, R.: Two limit theorems for random differential equations. Indiana Univ. Math. J. **22**, 1067–1089 (1972/73)
9. Coutin, L., Qian, Z.: Stochastic analysis, rough path analysis and fractional Brownian motions. Probab. Theory Relat. Fields **122**(1), 108–140 (2002). https://doi.org/10.1007/s004400100158
10. Dobrushin, R.L., Major, P.: Non-central limit theorems for nonlinear functionals of Gaussian fields. Z. Wahrsch. Verw. Gebiete **50**(1), 27–52 (1979). https://doi.org/10.1007/BF00535673
11. Dobrushin, R.L.: Gaussian and their subordinated self-similar random generalized fields. Ann. Probab. **7**(1), 1–28 (1979)
12. Deuschel, J.-D., Orenshtein, T., Perkowski, N.: Additive functionals as rough paths. Ann. Probab. **49**(3), 1450–1479 (2021). https://doi.org/10.1214/20-aop1488
13. Weinan, E.: Principles of Multiscale Modeling. Cambridge University Press, Cambridge (2011)
14. Friz, P., Gassiat, P., Lyons, T.: Physical Brownian motion in a magnetic field as a rough path. Trans. Am. Math. Soc. **367**(11), 7939–7955 (2015). https://doi.org/10.1090/S0002-9947-2015-06272-2
15. Friz, P.K., Hairer, M.: A Course on Rough Paths. With an Introduction to Regularity Structures, 2nd edn. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-41556-3
16. Fannjiang, A., Komorowski, T.: Fractional Brownian motions in a limit of turbulent transport. Ann. Appl. Probab. **10**(4), 1100–1120 (2000). https://doi.org/10.1214/aoap/1019487608
17. Fredlin, M.I.: Fluctuations in dynamical systems with averaging. Dokl. Akad. Nauk SSSR **226**(2), 273–276 (1976)
18. Friz, P., Victoir, N.: Differential equations driven by Gaussian signals. Ann. Inst. Henri Poincaré Probab. Stat. **46**(2), 369–413 (2010). https://doi.org/10.1214/09-AIHP202
19. Friz, P.K., Victoir, N.B.: Multidimensional Stochastic Processes as Rough Paths: Theory and Applications. Cambridge Studies in Advanced Mathematics, vol. 120. Cambridge University Press, Cambridge (2010). https://doi.org/10.1017/CBO9780511845079
20. Freidlin, M.I., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 3rd ed, vol. 260, xxviii+458. Springer, Heidelberg. Translated from the 1979 Russian original by Joseph Szücs (2012). https://dx.doi.org/10.1007/978-3-642-25847-3
21. Gonzales-Gargate, I.I., Ruffino, P.R.: An averaging principle for diffusions in foliated spaces. Ann. Probab. **44**(1), 567–588 (2016). https://doi.org/10.1214/14-AOP982
22. Gomez, C., Iyer, G., Le, H., Novikov, A.: An oscillator driven by algebraically decorrelating noise. arXiv:2107.07363 (2021)
23. Gehringer, J., Li, X.-M.: Homogenization with fractional random fields. arXiv:1911.12600 (2019)

24. Gehringer, J., Li, X.-M.: Diffusive and rough homogenisation in fractional noise field. arXiv:2006.11544 (2020)
25. Gehringer, J., Li, X.-M.: Functional limit theorems for the fractional Ornstein-Uhlenbeck process. J. Theor. Probab. (2021). https://doi.org/10.1007/s10959-020-01044-7
26. Gehringer, J., Li, X.-M., Sieber, J.: Functional limit theorems for volterra processes and applications to homogenization. Nonlinearity **35**, 1521 (2021)
27. Gyurkó, L.G.: Differential equations driven by Π-rough paths. Proc. Edinb. Math. Soc. (2) **59**(3), 741–758 (2016). https://doi.org/10.1017/S0013091515000474
28. Hairer, M.: An analyst's take on the BPHZ theorem. In: Computation and Combinatorics in Dynamics, Stochastics and Control. The Abel Symposium, vol. 13, pp. 429–476. Springer, Cham. arXiv:1704.08634 (2018)
29. Harris, T. E.: The existence of stationary measures for certain Markov processes. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II, pp. 113–124. University of California Press, Berkeley (1956)
30. Has'minskii, R.Z.: On the principle of averaging the Itô's stochastic differential equations. Kybernetika (Prague) **4**, 260–279 (1968)
31. Hairer, M., Li, X.-M.: Averaging dynamics driven by fractional Brownian motion. Ann. Probab. **48**(4), 1826–1860 (2020). https://doi.org/10.1214/19-AOP1408
32. Hairer, M., Mattingly, J.C.: Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations. Ann. Probab. **36**(6), 2050–2091 (2008). https://doi.org/10.1214/08-AOP392
33. Hairer, M., Majda, A.J.: A simple framework to justify linear response theory. Nonlinearity **23**(4), 909–922 (2010). https://doi.org/10.1088/0951-7715/23/4/008
34. Hairer, M., Mattingly, J.C.: Yet another look at Harris' ergodic theorem for Markov chains. In: Seminar on Stochastic Analysis, Random Fields and Applications VI. Progress in Probability, vol. 63, pp. 109–117. Birkhäuser/Springer Basel AG, Basel. arXiv:0810.2777 (2011). https://dx.doi.org/10.1007/978-3-0348-0021-1_7
35. Hairer, M., Mattingly, J.: The strong Feller property for singular stochastic PDEs. Ann. Inst. Henri Poincaré Probab. Stat. **54**(3), 1314–1340 (2018). https://doi.org/10.1214/17-AIHP840
36. Hairer, M., Mattingly, J.C., Scheutzow, M.: Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. Probab. Theory Relat. Fields **149**(1–2), 223–259 (2011). https://doi.org/10.1007/s00440-009-0250-6
37. Hairer, M., Schönbauer, P.: The support of singular stochastic partial differential equations. Forum Math. Pi (2022). https://doi.org/10.1017/fmp.2021.18
38. Jona-Lasinio, G.: Probabilistic approach to critical behavior. In: New Developments in Quantum Field Theory and Statistical Mechanics (Proceedings of Cargèse Summer Institute, Cargèse), pp. 419–446 (1976). NATO Adv. Study Inst. Ser. Ser. B Phys. 26 (1977)
39. Jacod, J., Shiryaev, A.N.: Limit theorems for stochastic processes. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, 2nd ed, xx+661. Springer, Berlin (2003). https://dx.doi.org/10.1007/978-3-662-05265-5
40. Kelly, D., Melbourne, I.: Deterministic homogenization for fast-slow systems with chaotic noise. J. Funct. Anal. **272**(10), 4063–4102 (2017). https://doi.org/10.1016/j.jfa.2017.01.015
41. Komorowski, T., Novikov, A., Ryzhik, L.: Homogenization driven by a fractional Brownian motion: the shear layer case. Multiscale Model. Simul. **12**(2), 440–457 (2014). https://doi.org/10.1137/13092068X
42. Kesten, H., Papanicolaou, G.C.: A limit theorem for turbulent diffusion. Commun. Math. Phys. **65**(2), 97–128 (1979)
43. Kurtz, T.G., Protter, P.: Weak limit theorems for stochastic integrals and stochastic differential equations. Ann. Probab. **19**(3), 1035–1070 (1991)
44. Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. Physica **7**, 284–304 (1940)
45. Kubo, R.: Stochastic Liouville equations. J. Math. Phys. **4**, 174–183 (1963). https://doi.org/10.1063/1.1703941
46. Kunita, H.: Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Mathematics, vol. 24. Cambridge University Press, Cambridge (1997). (**Reprint of the 1990 original**)
47. Kurtz, T.G.: Averaging for martingale problems and stochastic approximation. In: Applied Stochastic Analysis (New Brunswick, NJ, 1991). Lecture Notes in Control and Information Science, 177, pp. 186–209. Springer, Berlin (1992)
48. Ledesma, D.S., Borges da Silva, F.: Decomposition of stochastic flow and an averaging principle for slow perturbations. Dyn. Syst. **35**(4), 625–654 (2020). https://doi.org/10.1080/14689367.2020.1769031
49. Li, X.-M.: An averaging principle for a completely integrable stochastic Hamiltonian system. Nonlinearity **21**(4), 803–822 (2008). https://doi.org/10.1088/0951-7715/21/4/008
50. Li, X.-M.: Random perturbation to the geodesic equation. Ann. Probab. **44**(1), 544–566 (2016). https://doi.org/10.1214/14-AOP981

51. Li, X.-M., Sieber, J.: Mild stochastic sewing lemma, SPDE in random environment, and fractional averaging (2021)
52. Li, X.-M., Sieber, J.: Slow-fast systems with fractional environment and dynamics. Ann. Appl. Probab. arXiv:2012.01910 (2021) (**to appear**)
53. Lyons, T.J.: Differential equations driven by rough signals. Rev. Mat. Iberoam. **14**(2), 215–310 (1998). https://doi.org/10.4171/RMI/240
54. Maejima, M., Tudor, C.A.: Wiener integrals with respect to the Hermite process and a non-central limit theorem. Stoch. Anal. Appl. **25**(5), 1043–1056 (2007). https://doi.org/10.1080/07362990701540519
55. Meyn, S., Tweedie, R.L.: Markov Chains and Stochastic Stability, 2nd ed, xxviii+594. With a prologue by Peter W. Glynn. Cambridge University Press, Cambridge (2009). https://dx.doi.org/10.1017/CBO9780511626630
56. Mourrat, J.-C., Weber, H.: The dynamic Φ_3^4 model comes down from infinity. Commun. Math. Phys. **356**(3), 673–753 (2017). https://doi.org/10.1007/s00220-017-2997-4
57. Nualart, D.: The Malliavin Calculus and Related Topics. Probability and its Applications (New York), 2nd ed., xiv+382. Springer, Berlin (2006). https://dx.doi.org/10.1007/3-540-28329-3
58. Nualart, D., Xia, P.: On nonlinear rough paths. ALEA Lat. Am. J. Probab. Math. Stat. **17**(1), 545–587 (2020). https://doi.org/10.30757/alea.v17-22
59. Perruchaud, P.: Homogénéisation pour le mouvement Brownien cinétique et quelques résultats sur son noyau. PhD thesis, Université de Rennes 1 (2019)
60. Pei, B., Inahama, Y., Xu, Y.: Averaging principle for fast-slow system driven by mixed fractional Brownian rough path. J. Differ. Equ. **301**, 202–235 (2021). https://doi.org/10.1016/j.jde.2021.08.006
61. Papanicolaou, G.C., Kohler, W.: Asymptotic theory of mixing stochastic ordinary differential equations. Commun. Pure Appl. Math. **27**, 641–668 (1974). https://doi.org/10.1002/cpa.3160270503
62. Pipiras, V., Taqqu, M.S.: Integration questions related to fractional Brownian motion. Probab. Theory Relat. Fields **118**(2), 251–291 (2000). https://doi.org/10.1007/s440-000-8016-7
63. Rosenblatt, M.: A central limit theorem and a strong mixing condition. Proc. Natl. Acad. Sci. U.S.A. **42**(1), 43–47 (1956)
64. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, 3rd ed., xiv+602. Springer, Berlin (1999). https://dx.doi.org/10.1007/978-3-662-06400-9
65. Sinaı̆, J.G.: Self-similar probability distributions. Teor. Verojatnost. i Primenen. **21**(1), 63–80 (1976)
66. Skorokhod, A.V.: Asymptotic Methods in the Theory of Stochastic Differential Equations. Translations of Mathematical Monographs, vol. 78, xvi+339. American Mathematical Society, Providence. Translated from the Russian by H. H. McFaden (1989). https://doi.org/10.1090/mmono/078
67. Stratonovich, R.L.: A limit theorem for solutions of differential equations with random right-hand side. Theory Prob. Appl. **11**, 390–406 (1960). (**in Russian**)
68. Sethuraman, S., Varadhan, S.R.S., Yau, H.-T.: Diffusive limit of a tagged particle in asymmetric simple exclusion processes. Commun. Pure Appl. Math. **53**(8), 972–1006 (2000). https://doi.org/10.1002/1097-0312(200008)53:8<972::AID-CPA2>3.0.CO;2-
69. Tsatsoulis, P., Weber, H.: Spectral gap for the stochastic quantization equation on the 2-dimensional torus. Ann. Inst. Henri Poincaré Probab. Stat. **54**(3), 1204–1249 (2018). https://doi.org/10.1214/17-AIHP837
70. Veretennikov, A.Y.: On an averaging principle for systems of stochastic differential equations. Mat. Sb. **181**(2), 256–268 (1990)
71. Weinberg, S.: High-energy behavior in quantum field theory. Phys. Rev. (2) **118**, 838–849 (1960). https://doi.org/10.1103/PhysRev.118.838
72. Young, L.C.: An inequality of the Hölder type, connected with Stieltjes integration. Acta Math. **67**(1), 251–282 (1936). https://doi.org/10.1007/BF02401743